

A NEW MODEL FOR ROMANCE VERBAL CLITICS

by

CATHERINE SILLITOE

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of Italian Studies
(Modern Foreign Languages)
College of Arts and Law
University of Birmingham
September 2016

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Perlmutter (1971)'s seminal work on clitics has set much of the research model for ensuing studies. Despite enormous changes in linguistic theory over the intervening period, models in which clitic order is determined on the basis of grammatical person remains a key ingredient of most analyses. A key tenet of the current proposal is that clitic-forms may perform more than one syntactic function, reflected in their position within an elaborated series of feature projections including heads, not only for VP argument referents, but also non-argumental datives and nominative actors. Surface clitic patterns are merely sequential spell-outs of this structure. There is no need for clitic re-ordering at a morphological or syntactic level.

The proposed model requires no complex exclusion or conversion mechanisms, nor sophisticated syntactic processes, whilst being iconic and, therefore, learnable without the need for prior knowledge e.g. Universal Grammar constraints. The model has no need of lexicalized units, treating all clusters as purely compositional sequences directly interpretable from context. Giving each 'case' its own position leads to a simple and coherent model readily applicable across Romance. The work addresses 1-/2-/3-/4-clitic clusters in French, Italian, Spanish, Occitan, Catalan, and Romanian in their various dialect forms, whilst briefly illustrating many other Romance dialects.

Table of Contents

1 INTRODUCTION.....	1
1.1.1 Why Are Clitics Important?.....	1
1.1.2 Defining ‘Clitic’.....	2
1.1.3 Romance Clitics.....	4
1.2 Previous Approaches.....	5
1.2.1 Formalist Approaches.....	7
1.2.2 Application to Romance Clitics.....	9
1.2.3 Issues.....	11
1.3 Usage-Based Grammar.....	13
1.3.1 Grammaticalization.....	15
1.3.2 Lexicalization of Italian Clitics.....	18
1.4 A Communicative Approach.....	23
1.4.1 Explaining Exceptions.....	25
1.4.2 Non-Arbitrary Spanish Clitic-Clusters.....	27
1.5 Conclusions.....	30
2 MODEL.....	32
2.1 Elaboration of ‘Standard’ Models.....	32
2.1.1 The Current Model.....	36
2.1.2 Items Not Considered.....	38
2.1.3 Spell-Out.....	41
2.2 Against Reductionist Tendencies.....	45
2.2.1 Functions, <i>not</i> Forms.....	46
2.2.2 Syncretism.....	49
2.2.3 Null Entries.....	53
2.2.4 Unrealistic Expectations.....	55
2.3 Exclusions.....	57
2.3.1 RND.....	57
2.3.2 PCC.....	59
2.4 Conclusions to the Model.....	60
3 TWO DATIVES.....	62
3.1.1 OBL~DAT.....	62
3.1.2 [±E].....	63
3.1.3 Patterns Available.....	65
3.1.4 Chapter Outline.....	67
3.2 Lower Clitic-Field.....	68
3.2.1 A Note on Translations.....	68
3.2.2 ‘Dative’ ≠ Possession.....	69
3.2.3 ‘Dative’=Affectedness.....	70
3.2.4 (In)alienable Possession.....	73
3.2.5 Clitic Doubling.....	74
3.2.6 Conclusions for the Lower Clitic-Field.....	75
3.3 Upper Clitic-Field.....	75
3.3.1 Sympathetic.....	76

3.3.2 Settings.....	80
3.3.3 State, <i>not</i> Place.....	82
3.3.4 Possession.....	85
3.3.5 Restrictions.....	87
3.3.6 Inferences of Causation.....	89
3.4 Separating Fields.....	92
3.4.1 Absence of OBL _[+R]	92
3.4.2 Laísta Dialects.....	93
3.4.3 Lower Benefactives.....	94
3.4.4 Emphatics.....	98
3.4.5 Putative PCC-Breaches.....	99
3.4.6 Conclusions.....	101
3.5 Communication Theory and Clitic Patterns.....	102
3.5.1 Signalling Relationships.....	102
3.5.2 Parsing and Efficiency of Communication.....	104
3.6 Conclusions.....	107
4 THE UBIQUITY OF SE.....	108
4.1 Introduction.....	108
4.1.1 The Problem.....	109
4.1.2 Unity vs. Diversity.....	110
4.2 Reflexive SE.....	112
4.2.1 Reflexive Functions.....	112
4.2.2 Contrastive Pronominals.....	115
4.2.3 Case.....	117
4.2.4 Emphatics.....	118
4.2.5 Reflexives ≠ Intransitive.....	121
4.2.6 Anticausatives ≠ Reflexives.....	124
4.2.7 Conclusions for Reflexivity.....	128
4.3 Non-Reflexive SE.....	128
4.3.1 Morphological Marking.....	130
4.3.2 Variations.....	132
4.3.3 Restrictions on Application.....	136
4.3.4 Proposal.....	140
4.3.5 Properties.....	142
4.3.6 Outline.....	145
4.4 Non-Actives as a Class.....	145
4.4.1 Event-Passives.....	146
4.4.2 Control.....	148
4.4.3 ‘Agentive’ Adverbs.....	151
4.4.4 <i>By</i> -Phrases.....	154
4.4.5 Other Prepositions.....	156
4.5 Non-Actives as a Mechanism.....	158
4.5.1 Romance Development.....	160
4.5.2 Non-Actives in Contrast.....	163
4.5.3 Derivation.....	164
4.6 SE _{PASS} ~SE _{IMP}	168

4.6.1 The Constructions.....	168
4.6.2 SE _{PASS}	170
4.6.3 SE _{IMP}	171
4.6.4 SE _{PASS} ≠ SE _{IMP}	172
4.6.5 SE _{NAR}	174
4.6.6 SE _{PASS} >SE _{NAR} /SE _{IMP}	178
4.6.7 Non-Concordance.....	180
4.6.8 Spanish SE _{IMP}	183
4.6.9 Italian SE _{IMP}	184
4.6.10 Other Variations.....	188
4.6.11 Exclusions and Substitutions.....	190
4.7 SE _{ANT} ~SE _{NOM}	192
4.7.1 SE ≠ Dative.....	196
4.7.2 SE _{NOM}	198
4.7.3 SE _{ANT}	202
4.7.4 Verbs of Motion.....	204
4.7.5 ‘Pronominal Verbs’.....	206
4.7.6 Putative Metathesis.....	210
4.8 Composition and Interpretation.....	211
4.8.1 Conclusions for SE.....	213
4.8.2 Adequacy of Form(s).....	214
4.8.3 Adequacy of Model.....	216
4.9 Conclusions.....	216
5 NON-PERSONAL CLITICS.....	218
5.1.1 Against Lexicalization.....	218
5.1.2 Interpretation.....	220
5.1.3 Range/Categories.....	222
5.1.4 Forms.....	224
5.1.5 Chapter Outline.....	224
5.2 Object-Oriented Clitics.....	225
5.2.1 <i>Ci</i>	225
5.2.2 <i>Ne</i>	228
5.2.3 Object-Clitic/Functions.....	231
5.3 Subject-Oriented <i>Ne</i>	232
5.3.1 <i>Ne</i> _{NOM} ~ <i>Ne</i> _{OBL}	235
5.3.2 <i>Ne</i> _{ABL}	240
5.4 Subject-Oriented <i>Ci</i>	240
5.4.1 Existentials.....	240
5.4.2 Romance Existentials.....	242
5.4.3 Italian.....	245
5.4.4 Romanian.....	250
5.4.5 Sardinian.....	251
5.4.6 Diversity of <i>Ci</i> _{EXI}	253
5.4.7 Exclusions.....	255
5.4.8 Conclusions.....	256
5.5 Putative ‘Lexicalization’.....	256

5.5.1	LOPHRASAL/LaABSTRACT.....	257
5.5.2	Se+Lo/La.....	259
5.5.3	Object-Oriented Ce+La.....	261
5.5.4	Object-Oriented Ne.....	265
5.5.5	(Ci)+Se+Ne.....	271
5.5.6	Subject-Oriented Ne _{ABL}	274
5.5.7	Subject-Oriented Ci.....	279
5.5.8	Summary.....	283
5.6	Conclusions.....	284
6	SWAPPING.....	286
6.1	Introduction to Swapping.....	286
6.1.1	The D/A~A/D Parameter.....	287
6.1.2	Spell-Out.....	289
6.1.3	Chapter Outline.....	290
6.2	The Nature of Spurious 3-3.....	290
6.2.1	Orthography and Structure.....	290
6.2.2	Morphemic Structure and Markedness.....	291
6.2.3	3-3-Rules.....	292
6.2.4	Motivation/Nature of OTHER.....	293
6.2.5	Development of <i>Gli</i>	294
6.2.6	Generalisation of <i>Gli</i>	296
6.2.7	3-3-Rules Across Romance.....	297
6.3	Italian.....	300
6.3.1	Basic Patterns.....	300
6.3.2	Prosody.....	303
6.3.3	Locatives.....	305
6.3.4	Syntactic Approaches?.....	306
6.4	Catalan.....	307
6.4.1	Sequence-Variation.....	309
6.4.2	Complex Forms.....	312
6.4.3	[(ə)lz(ə)ni]/[(ə)lzin].....	315
6.5	Occitan.....	317
6.5.1	Development.....	318
6.5.2	Provençal.....	321
6.5.3	Languedocian.....	322
6.5.4	Gascon.....	323
6.6	Aragonese.....	326
6.7	Proclisis: Conclusions.....	330
6.8	Enclisis.....	331
6.8.1	WP Status.....	331
6.8.2	L-Allomorphs & Sequence.....	333
6.8.3	L-Allomorphs & Displacement.....	335
6.8.4	Prosodic Structure.....	337
6.8.5	Verb PW Boundary.....	339
6.8.6	Lucanian.....	342
6.8.7	Neapolitan.....	344

6.8.8 Sardinian.....	345
6.8.9 Sardinian II.....	347
6.8.10 Catalan.....	348
6.9 Conclusions for Enclisis.....	350
6.10 French.....	352
6.10.1 Prosodic Structure.....	352
6.10.2 Against WPs.....	354
6.10.3 Development.....	356
6.10.4 Analysis.....	360
6.11 3-3-Rules.....	363
6.11.1 Putative Feature Transfer.....	364
6.12 Weight.....	368
6.12.1 Conclusions.....	369
7 EXCLUSIONS.....	370
7.1 Introduction.....	370
7.2 Proposition.....	371
7.3 Person-Ordering.....	373
7.3.1 Person/Number Restrictions.....	376
7.4 Present-Verbs.....	379
7.4.1 Cases.....	382
7.4.2 Constructions.....	384
7.4.3 Western Romance.....	387
7.4.4 Romanian.....	391
7.4.5 No PCC-Violations.....	399
7.4.6 Old Spanish.....	401
7.4.7 PCC Conclusions.....	404
7.5 Exclusions.....	404
7.5.1 4-Clitic Clusters.....	404
7.5.2 Function, <i>not</i> Form.....	405
7.5.3 Delimiting the PCC.....	406
7.5.4 Proscriptions.....	408
7.5.5 Re-Evaluating RND/PCC.....	410
7.6 Conclusions.....	415
8 CONCLUSIONS.....	416
8.1 Summary.....	416
8.2 Areas Not Covered.....	423
8.3 Conclusions.....	425
9 CORPORA.....	428
10 BIBLIOGRAPHY.....	430

LIST OF ABBREVIATIONS

[±ANIM]	Animacy
[±DEF]	Definiteness
[±E]	Externality
[±R]	Reflexivity
[±SPEC]	Specificity
ABL	Ablative
ACC	(A) Accusative
AUX	Auxiliary Verb
CG	Clitic-Group
CL	Clitic
COS	Change of State
CP	Complement Phrase
DAT	(D) Dative
DOC	Double Object Construction
DOM	Double Object Marking
DP	Determiner Phrase
EXI	Existential
GEN	Genitive
IMP	(I) Impersonal
IP	Inflectional Phrase
LDA	Long Distance Agreement
LOC	Locative
MC	Morphological Component
NEUT	Neuter
NOM	(N) Nominative
NP	Noun Phrase
OBL	(O) Oblique
PCC	Person Combination Constraint
PolP	Polarity Phrase
PP	Prepositional Phrase
PPh	Prosodic Phrase
PRT	Partitive
PW	Prosodic word
REFL	Reflexive
RND	Referent Non-Duplication
SCL	Subject Clitic as found in Northern Italy/Gallo-Romance
SE _{ACC}	Any Reflexive Clitic in Accusative Position
SE _{ANT}	Any Reflexive Clitic used in Anticausative Constructions
SE _{DAT}	Any Reflexive Clitic in Dative Position
SE _{IMP}	Any Impersonal Nominative Clitic used in Impersonal Constructions
SE _{MID}	Any Reflexive Clitic used in Middle Constructions
SE _{NAR}	Any Impersonal Nominative Clitic used in Existential Constructions
SE _{NOM}	Any Reflexive Clitic in Nominative Position

SE _{PASS}	Any Reflexive Clitic used in Passive Constructions
SE _{SPUR}	Spurious Replacement Clitic in Spanish e.g. <i>le+lo</i> → <i>se_{SPUR}+lo</i> .
S _H	High Subject Position preceding the Verb
S _L	Low Subject Position following the Verb
SOA	State of Affairs
SUBJ	Subject
VP	Verb Phrase
WP	Weak Pronoun

Languages

[CA]	Catalan
[FR]	French
[IT]	Italian
[RO]	Romanian
[SP]	Spanish

1 INTRODUCTION

Perlmutter (1971)'s seminal work on Spanish clitics has set much of the research model for ensuing studies. Despite enormous changes in linguistic theory over the intervening period, a model in which clitics are ordered on the basis of constraints/mechanisms centred on grammatical person (person-models, §1.2.2) remains a key ingredient of most analyses. This work provides a model based on case (case-model, Chapter 2) which provides simpler and more comprehensive results.

1.1.1 Why Are Clitics Important?

In Romance, whilst new or (re-)topicalized verbal arguments are expressed as full DPs (a), arguments already in discourse are represented by clitics (b). Such clitics (usually monosyllabic) substitute a range of arguments requiring whole phrases (3-4), or having no equivalent (5-7) in English, whilst re-using single forms for multiple functions.

Table 1

	(a)	(b)	
1	Mando una carta .	La mando.	I send a letter/it .
2	Mando una carta a Maria .	Le mando una carta.	I send a letter to Maria/her .
3	Mando una carta a Maria .	Gliela mando.	I send a letter/it to Maria/her .
4	Mando una carta a Roma .	Ci mando una carta.	I send a letter to Rome/there .
5	'Anticausative'	I piatti si rompono.	The plates break.
6	'Passive'	I libri si vendono qui.	Books are sold here.
7	'Impersonal'	Si mangia bene qui.	One eats well here.

Clitics may combine (3b) but are subject to complex combinatorial constraints and mutations, for which a single coherent model has proved illusive.¹ Explanation of clitic systems is fundamental to any theory of communication as their anaphoric properties act as the glue

¹ "for more than a quarter century, French pronominal affixes...have posed a dilemma for generative grammar" (Miller & Sag 1997:573).

which enables separate utterances to become meaningful and efficient discourse, tracking significant actors/objects across sentences, and expressing the same message from various perspectives (e.g. active, passive, middle) with little or no change in the rest of the sentence.

The fascination of clitics revolves around how interlocutors can compose, interpret, and re-compose shared views of situations through infinitesimally small amounts of data, which appear at first sight to be entirely inadequate to carry such a huge burden of meaning, and too limited in form(s) to allow distinction between their manifold uses. Moreover, we want to understand the source of the restrictions which are so often treated as arbitrary.

Clitics offer a window into the details of verbal structure and how meaning is composed and parsed. This work presents a model where technical details arise naturally from semantic and syntactic structure, which when combined with devices such as focus and topicalization within broader pragmatic contexts results in a situation where clitics are optimally suited for their task, and their behaviour is fully predicted, rather than exotic.

1.1.2 Defining ‘Clitic’

Zwicky (1977) defines *clitics* as (a) phonological simplifications of full forms which attach phonologically to hosts e.g. English *bring 'em*; (b) simple clitics which are not reduced forms but must lean on another word in order to be prosodically realized e.g. Latin *Senatus Populus-que Romanus*; (c) special clitics such as Spanish *se*, which have developed specialised morpho-syntactic behaviour. These classes are not closed, with (a)/(b) often developing into (c) over time. They form a heterogeneous category including pronouns, auxiliaries, determiners, negative particles, and interrogative particles (Klavans 1982, 1985; Riemsdijk 1999; *i.a.*).

Romance clitics are variously described: clitics, morphemes, affixes, often with particular functionality e.g. SE as valency reducing operator (Baauw & Delfitto 2005:165). Fontana (1994)'s historical and dialectal study of Spanish, proposes that whilst clitics were once pronouns, they have become morphemic; diverse dialectal behaviours being evidence of developmental stages. Franco (1993) considers that we are in the midst of evolution from pronominals to affixes.

Putative proofs of morphemic status include (1) exclusive hosting by verbs, but this was not true in earlier times, and even today e.g. Italian *ecco+lo*; (2) clitics form rigid orders like morphemes and unlike words, but this does not argue for morpheme status but rather against independent word status; (3) some dialects allow interchange of 3.PL desinence and clitic e.g. *márche+se+n~márche+n+se* (Oroz 1966:310), however, only this desinence is involved and it would be as reasonable to argue that *n(o)* which came to be added to 3.PL in order to differentiate it from 3.SG remains an independent unit in these dialects as in earlier stages of Romance (Maiden 1995); (4) clitics and morphemes are unaccented (Fernández Soriano 1999:1252), however, while clitics do not bring their own stress, it is common in speech to find those following imperatives carrying main verbal accent. Alvar & Pottier (1983:§98) even note their graphic marking in Golden Age poetry (*Comportesé*). This is hardly unequivocal evidence.

Otero (1999:1472, 2002:168-71) notes that SE has properties found in no verbal morpheme e.g. appearing as enclitic (*Afēita+te*), proclitic (*Pedro se afeitó*), and far from principal verbs when auxiliaries are present (*Juan se quiso afeitar*), however, Franco (2000:182) provides cross-linguistic examples of verbal morphemes equally separated from their verbs,

considering such behaviour to be a natural possibility of morphemes. The definition of *morpheme*, therefore, appears to be as loose as that of *clitic*, and indeed those who favour morphemic analyses, use the same arguments to arrive at different results, considering all clitics to be morphemes, or just reflexives, or just *se*. The *morpheme~clitic* debate adds little to our understanding.

This work focuses upon *what appears where and when in the surface form*, since this is what listeners must parse for communication to occur. From this perspective, *morpheme* or *affix* are simply labels which because of use in other fields bring with them connotations which are often inappropriate to this area of investigation. Indeed, Zwicky (1994:xiii) considers *clitic* as “an umbrella term, not a genuine category in grammatical theory”. Similarly, Sadock (1995:260) claims “there is [not] a natural class of clitics defined in terms of genuine grammatical properties...[T]he various things which have been put in this category by linguistic researchers do have something sociological in common, namely their reluctance to fit naturally into any single one of the classical components that traditional grammar recognizes.” We follow Fernández Soriano (1999:1251)’s advice to use the term *clitic* exactly because it lacks any clear definition beyond that given by Zwicky.

1.1.3 Romance Clitics

Modern Romance clitics (henceforth, simply *clitics*) developed through phonological weakening from Latin personal pronouns and locative adverbials. Initially clitics attached to any host, subject to the Tobler-Mussafia Law which precluded clause-initial position. Relics survive e.g. with expletives (Italian *ecco+lo*, Romanian *iată-l*, ‘here it is’), certain prepositions in Old Italian (*(in)contro/allato+gli* ‘against/beside him’, and some modern Northern Italian dialects (Renzi 1988:359, fn.12). From XIII^c (Maiden 1995), clitics became

increasingly centred upon the verb irrespective of clausal position, and fixed in their order relative to each other. Their position relative to the verb varies cross-linguistically:

Table 2

	Finite	Infinitive	Participle/Gerund	Imperative
Spanish/Italian	pre-verbal	post-verbal		
Romanian	pre-verbal		post-verbal	
French	pre-verbal			post-verbal

Clitics are often represented as highly idiosyncratic. Viewed from traditional perspectives, clitics appear to present combinatorial restrictions, re-ordering, and opaque forms, which are often labelled by means of an example. Putative restrictions and means of enforcement are wide and varied. We hope to show that the situation is, in fact, quite simple when viewed from case, rather than person.

Table 3

Exclusions	French	* <i>me+lui</i>	1/2-person pronouns may not precede <i>lui</i> .
	Spanish	* <i>me+se</i>	No personal pronouns may precede <i>se</i> .
	General	* <i>me+te</i>	No 1+2 or 2+1 combinations.
Swapping	French	* <i>lui+le</i>	<i>lui+le</i> → <i>le+lui</i> .
Opaque	Spanish	spurious- <i>se</i>	<i>le(s)+lo/a(s)</i> → <i>se+lo/a(s)</i> .

Most of this work focuses upon proclitic order, which displays the most complex patterns. Chapter 6 explores post-imperative sequence variations which follow from the same model.

1.2 Previous Approaches

This section reviews various perspectives available for modelling grammars, in relation to syntactic variation vs. exceptions and ungrammaticality, with particular reference to clitics.

The central issue, in our opinion, is willingness to accept arbitrariness of language (as preferred explanation), in general, and in particular with reference to ‘anomalous’ clitic

behaviour. This is ‘reasonable’ from the formalist view point and its notorious *autosyn* hypothesis (§1.2.1), but leads to issues being prematurely exiled to morpho-prosody, attributed to “weird morphological constraints” (Bonet 1994:51), no longer part of syntax or even semantics: “[c]litic clustering is...a matter of considerable irrelevance to pure formal syntax...it almost does not impinge on it” (Wanner 1994:51, my translation).

Ironically, usage-based grammars (§1.3) which repudiate *autosyn*, end up creating new ways to accept arbitrariness through reliance upon lexicalized (i.e. stored and, therefore, non-analysable) words/phrases: “from the assumption that the lexicon is the repository of irregularity, many lexicalists seemed to derive the conclusion that language is one great trove of irregularity” (Newmeyer 1998:219). Whilst each approach provides valuable insights, ultimately, they leave language as random collections of disconnected items, rather than something organic, interpretable, and usable as means of communication. They deny/ignore the compositional and interpretive dimension of language.

§1.4 considers cognitive/communicative perspectives which stress language’s essential iconicity, acquisition through communication, and variation’s *positive* role in syntactic analysis. Acceptability variation and exceptions are seen in terms of cognitive processes of interpretation of messages within context, without recourse to arbitrary removal of non-analysable chunks. §1.4.2 considers García (2009)’s study of Spanish clitics which aims to show that frequency of variations and exceptions are motivated by cost of cognitive analysis. It bases its analysis on mapping semantics directly to surface sequences, implicitly following Manning (2003:313)’s denial of our ability to determine underlying structure. Whilst providing considerable insight into negative exceptions, extension to the constraints on

combinations of personal clitics (PCC) is, in our opinion, less successful, failing to distinguish variably acceptable variations and ‘negative’ exceptions from perfectly reasonable but aberrantly unacceptable ‘positive’ exceptions. In order to show flaws in *autosyn*, García creates a model unrelated to (i.e. autonomous from) structure. Formalism focuses on structure ignoring meaning, whilst García’s strong functionalist view focuses on meaning ignoring structure. Ultimately, neither is successful.

This work presents a model drawing insights from all these approaches, which not only takes account of structure but explains ‘positive’ exceptions in terms of that structure. It retains interpretation as the explanation of ‘negative’ exceptions and indeed the driving force behind why structure is as it is. This structure allows interpretation *in context* of any combination by composing meaning from its constituent parts, thereby removing the need for arbitrary rules or lexicalization, and bringing clitics back into the heart of syntax.

1.2.1 Formalist Approaches

Whilst “syntax involves the stringing together of independent sub-units into a longer signal” (Hurford 2003:43), allowing infinite numbers of complex signals, not all sequences are equally acceptable. Beyond social/normative control, this property is generally referred to as *grammaticality*, which (Chomsky 1957:16 *et pass.*) considers to be of prime importance (independent of meaning or frequency of use), presupposing that the set of grammatically well-formed sentences is “somehow given in advance” (Chomsky 1957:85), and may be identified “on the basis of context-isolated acceptability judgements” (Newmeyer 1998:59). The formalist approach posits rules and structures to generate this set, independently of meaning (Stefanowitsch 2007:62), opposing itself to the common view that sequences are (un-)grammatical only “under the intended interpretation” (Stepanov *et al.* 2004:79).

The *autosyn hypothesis* (Newmeyer 1998:28) defines syntax as autonomous, involving three tenets: (a) some elements of syntax are arbitrary (arbitrariness); (b) arbitrary elements participate in systems (systematicity); (c) systems are self-contained (self-containedness). As often noted (e.g. Matthews 1979:210-13; Schutze 1996:29-30; Wasow & Arnold 2005), this makes *autosyn* and *grammaticality* circularly interdependent and self-fulfilling. ‘Exceptions’ become seen as mere grammatical vagaries rather than counter-evidence for arguments, or even *prima facie* evidence for arbitrariness in autonomous syntax (Hudson *et al.* 1996).

Whilst *positive exceptions* are items/arrangements which should not undergo rules but do, *negative exceptions* are cases which fail to undergo rules for which they are eligible. In either case, the predicted ‘grammatical’ output fails to be observed and is considered unacceptable. A classic case of arbitrariness resulting in negative exceptions is the English “double-object dative”, for “there are verbs that fit the semantics of the dative but cannot use it [sic], ...Tell/*Explain Bill the answer” (Jackendoff 1997:175). This creates a central problem for language acquisition; Baker’s Paradox, or how children can learn to avoid plausible yet unacceptable combinations, given that non-occurrences cannot be observed (cf. Fodor 2001:369-70; Stefanowitsch 2008).

Pinker (1989) attempts to reconcile Baker’s paradox within formalist treatments, by pushing difficulties into the lexicon, such that each surface variation is a separate lexical entry with “property-predicting” linking-rules mapping them onto particular surface forms (p.71-72) and semantically to each other (p.94-5). The ultimate conclusion seems to be that *throw* dativizes, but *pull* does not, because only the former implies a receiver within the event, matching prepositional forms.² Unfortunately, “[w]e currently have neither a format for the input

2 It is acceptable in requests to barmaids to “pull me a pint”, which use benefactive rather than goal datives.

structure of a rule nor a matching function by which a semantic structure for a word would be deemed to match or not to match a rule” (p.213). The results are unconvincing, and often self-contradictory. García (2009) for a detailed critique.

Defining **explain Bill the answer* as a negative exception to a lexical/syntactic rule implies equivalence to the learning of lexically idiosyncratic morphological irregularities (Bowerman 1988, 1996; Roberts *et al.* 2005:334); indeed, Jackendoff (2002:191) claims that “marked rules deviate from the unmarked case qualitatively in just the way irregular verbs deviate from regular forms.” However, the two sets of irregularities are not comparable: while it is possible to list English irregular plurals, this is impossible for English double-object structures (Aissen & Bresnan 2004:581); over-generalization is common with morphological patterns but rare in syntax (Howell & Howell 2006:882); pre-emptive blocking of an ‘ungrammatical’ generalisation is operative in the learning of inflections, but not syntax (Braine & Brooks 1995:359-60), where ‘correct’ usage may coexist for years with syntactic over-generalization (Bowerman 1996:461-3). The only way that formalists can deal with such irregularities is to exile them from syntax i.e. ignore them.

1.2.2 Application to Romance Clitics

Since clitics exhibit numerous positive and negative exceptions, accounting for impossible clusters is relegated to functional dimensions external to formal grammar (Wanner 1994:30) or assigned to autonomous morphological components (henceforth MC, e.g. Bonet 1995a; Harris 1996, 1997).

This implies that each verb would require several separate entries.

Perlmutter (1971:38) argues that templates are required to generalise ordering and exclusion of clitics because some “well-formed deep structures correspond to no grammatical surface structure. Only a surface constraint can characterise such sentences as ungrammatical.” For Spanish, “clitics are *strictly ordered*” (p.46, original italics) as *se>II>I>III*. Grammatical sequences are defined in terms of person rather than grammatical function, whilst combinations are excluded based on surface form alone.³

Subsequent debate concerning the theoretical status of templates has proved fruitless (e.g. Dinnsen 1972; Wanner 1994). With no principled theory, templates remain unconstrained devices added to morpho-syntactic derivations without any motivation other than to *describe* attested but still unexplained facts. Problems have long been apparent; Wanner (1977) notes that not all clitics respect transitivity as required by templates, while Strozer (1976:171) notes that templates require rules referencing function normally disallowed in surface constraint models. Harris (1996) notes that a four slot template creates the unfulfilled expectation that four clitic-clusters will be as likely as smaller clusters, while Cuervo (2003) notes that, since competition for slots is symmetric, a template cannot choose between two clitics. Such underlying problems are reflected in practical flaws; templates often ban grammatical structures while accepting ungrammatical ones.

Alternative mechanisms (but with an identical target) using syntactic movement suffer from the problem: movements should be controlled by source position/function, but template targets are controlled by person (Heap & Roberge 2001 for an overview). Solutions (e.g. Bastida 1976; Uriagereka 1995) which distinguish 3-person from 1/2-person clitics based on some positional difference in syntactic heads, do so by introducing syntactic movements

³ In this work, the term ‘person-model’ is used to cover the numerous variations upon this approach.

which are entirely unmotivated other than to *describe* these surface orders. Accounts based upon ‘base generation’ fair little better. Bonet (1991, 1994, 1995a, 1995b) employs an MC able to manipulate clitic morphological structure, but provides no principled account of why featural content might determine a clitic’s position relative to another. Harris (1994, 1996) proposes “precedence conditions” which constrain ordering relationships between different (groups of) clitics, whilst optimality approaches (e.g. Anderson 1996; Grimshaw 1997) use ALIGN constraints to place clitics in relationship to each other. In all these approaches, conditions/constraints are unmotivated other than to *describe* the apparent ordering facts. The methods are *ad hoc*, un-generalizable and non-predictive. Even if it were possible to modify such proposals in order to satisfy all the data, it would add nothing to our understanding; simply exchanging one set of unmotivated proposals for another.

1.2.3 Issues

Empirical studies show that many clitic-clusters do not conform to person-ordering and the basis of this condition is an excessive idealisation of the data: Perlmutter (1971:50-51) notes dialect variation in 2-clitic sequences; Bastida (1976) itemises even greater variation for 3-/4-clitics clusters; whilst the **me+se* restriction is so commonly broken that it requires specific prohibition in the standard’s official grammar.⁴ Such non-compliant data is ‘left for future research’ or partially handled by adding increasingly complex structures and/or processes to force recalcitrant clitics into their idealised position. The goal of person-ordering is derived from an unrepresentative data sample and should not guide our investigations. This work attempts to deal with the whole data set.

4 RAE (1973:427) considers it “solecismo plebeyo”, however, it has featured in Spanish (Martín Zorraquino 1979:347-352) and other Romance varieties (Hetzron 1977) for centuries.

Each clitic surface-form is treated identically regardless of its contextual semantic/syntactic function, however, Romance's development has seen many shifts of form and function. Italian *ci/vi* replaced *nos/vos* to become 1/2.PL personal clitics, whilst retaining their locative value in other contexts (§5.2.1). Precedence of function over form is illustrated by the French **me+lui* constraint which applies to indirect-object, but not ethical, datives despite identical forms (Kayne 1994). Analyses are often inconsistent. Whilst Italian *ci=we* and *ci=here* are distinguished despite identical forms, French *y* is treated as a unity despite its separate functions being easily distinguished by syntactic behaviour. As Heggie & Ordóñez (2005:12-13) show, apparent ordering conflicts of *y* evaporate when these are taken into account.

Autosyn's exiling of clitics from syntax leads to consideration of clitic-clusters in isolation from the grammar of which they are but a small part. Everything is expressed in terms of exclusions/orderings of clitic forms *in vacuo* rather than the arguments which they express. This leads to rules banning sequences because they are unacceptable in one context even though they are legitimate in others. The accumulation of such context-free rules makes it impossible to deal with, or even worse make false predictions about, larger sequences. This work starts from the premise that by considering the function of each clitic in context, it is possible to see why *particular sequences* are unacceptable in *particular situations*.

In following chapters, we hope to show that focus on surface-forms combined with adherence to *autosyn*, and thereby premature acceptance of arbitrariness, has lead to functionally distinct impersonal, passive, transitive, and spurious-*se* being lumped together (Chapter 4),⁵ even though they are semantically, syntactically and logically mutually exclusive. Similarly, the implications of two types of datives with different syntactic properties *and* position (Chapter

5 Grimshaw (1997) considers *se* a default form surfacing whenever constraints ban everything else.

3), and the very existence of nominative clitics, have been ignored simply because they have identical surface forms. The overall effect is to ‘smudge’ surface forms across syntactic positions making templates and mapping *appear* necessary. If, however, clitics are given their appropriate place in case-oriented models, they always appear in sequences determined by semantic function, matching that of the final syntax tree, requiring no mapping within a MC, and (almost) no exclusions.

1.3 Usage-Based Grammar

Autosyn is rejected by those who see grammars as emerging from use, as successive generations of learners abduct competence/langue from performance/parole (Bybee & Hopper 2001; Hurford 2003:54; Kirby & Brighton 2004:592; *i.a.*). Whilst the quantitative aspects of language are irrelevant to supporters of *autosyn* (Scholz & Pullum 2007:715), they are central to usage-based grammars.

Frequency of use is implicated in language acquisition removing the need for innate Universal Grammar/language acquisition devices (Redington *et al.* 1998; Rohde & Plaut 1999:105; Marcus 1999; Culicover 1999:197; Mintz *et al.* 2002; Tomasello 1995, 2000), whilst diachronically, entrenchment of frequent collocations favours categorical recourse to them, creating non-analysable units (Lüdtke 1980; Bybee 2006:714-16). Whilst informative in fields with finite numbers of discrete units (Bybee 2001; Pierrehumbert 2003), its value is less evident in morphology (Pinker 1998) and questionable in open-ended syntax (Newmeyer 2003, 2005), where it is impossible to identify finite sets of types (Sampson 2001:170-178; Goldberg 2002:340-41; Hawkins 2004:16).

In frequency-based analyses, acceptability is relative (dependent on intended reading), and ‘non-occurrence’ is simply an extreme decrease in frequency (approaching zero), relative to competing options. Whilst event frequency appears to be automatically encoded in the brain (Hasher & Zacks 1984; Manning 2003), it does not follow that internalised probabilities account for greater frequency of particular items in actual language use (Wasow & Arnold 2003:133; Bresnan 2006), since quantitative skewing imputed to internalised lexical biases might merely reflect grammar-external “performance” factors (Kiparsky 1971:603). Indeed, Green (2004:330) considers “arbitrary lexical bias...is not so much an explanatory factor as...an effect in search of an explanation.”

Construction Grammar is characterised by focus upon frequency-based internalisation of complex units (Croft & Cruse 2004:155; Culicover 1999:33; Sag & Wasow 1999:369; Wray 2002:15; Stefanowitsch & Gries 2003:209-11). Such *Constructions* are defined as form-meaning pairs, where some aspect is not strictly predictable from its component parts (Goldberg 1995:4), making them symbolic units, comparable to conventional lexical signs (Kay 1997:123; Langacker 2005:140-43; Croft & Cruse 2004:247; Stefanowitsch & Gries 2003:209-11). However, postulating that frequent sign-combinations are automatically internalised as *Constructions* ignores the compositionality of utterances and avoids discussion of both units and calculus (Bybee & Eddington 2006:328). If frequency alone determines constructional status, retrieval of these “preferred strings” becomes indistinguishable from their preferential composition in response to frequent communicative needs (Wray 2002:7). Before a construction can be attributed independent status, it must be determined whether its meaning can be “computed from the meanings of the individual words and the way they are

arranged” (Pinker 1998:220), however, this kind of demonstration is rare in constructionalist discussions. Moreover, the focus of usage-based grammar upon *lexicalization* of highly frequent collocations leaves it with little to say about syntactic productivity, or (un)acceptability of daily extrapolations from the norm (Barlow 2000).

1.3.1 Grammaticalization


Grammaticalization has diverse interpretations (Bisang *et al.* 2004), but essentially describes a broad diachronic process where forms lose syntactic independence becoming increasingly grammatically circumscribed. Considered epiphenomenal by some (“nothing more than a label for the conjunction of certain types of independently occurring linguistic changes”, Newmeyer 1998:237), it remains a useful “research framework” (Hopper & Traugott 2003:1), representing “the most salient case of a pervasive regularity of language change” (Haspelmath 2000:248). For clitics, it is the discourse/pragmatic phenomena of language change (unavailable to formalists) which provides key evidence for understanding their synchronic and diachronic behaviour.

Cross-linguistic studies show that these shared processes tend to follow similar patterns,⁶ favouring particular lexical classes: frequently used terms become more abstract (Latin HOMO ‘man’>French impersonal *on*); demonstrative pronouns lose their deictic meaning evolving into definite articles (Latin ILLE ‘that’>French *le* ‘the_{M.SG}’). Loss of syntactic autonomy is generally accompanied by reduction in phonetic/phonological status (*phonetic erosion*) and semantic substance (*semantic bleaching*). A complex example is development of

6 Cross-linguistically, grammaticalization strongly favours suffixation over prefixation. Klausenburger (2000) proposes that the crucial role of initial words and/or segments for perception makes them less likely to undergo more advanced stages of grammaticalization to produce prefixes.

Latin analytic constructions (infinitives+present/perfect auxiliary), where independent auxiliaries became bound morphemes i.e. inflectional desinences of synthetic future/conditional tenses (8,Vincent & Harris 1982; Klausenburger 2000; Schwegler 1990; *i.a.*).

Table 4

8	Latin		Italian	
	CANTARE _{INFINITIVE} HABET _{PRESENT.TENSE}		canter-à _{FUTURE}	S/he will sing
	CANTARE _{INFINITIVE} HABUIT _{PRESENT.PERFECT.TENSE}		canter-ebbe _{CONDITIONAL}	S/he would sing

Grammaticalization is multi-dimensional, occurring along various continua⁷ expressed across different aspects of grammar, not necessarily reaching completion in any dimension. Such continua are not ordered sequences of discrete units, but overlapping phases allowing transition over time. Synchronically, it expresses the range of alternatives available to realize linguistic construals, and is “primarily...a syntactic, discourse pragmatic phenomenon, to be studied from the point of view of fluid patterns of language use” (Hopper & Traugott 2003:2). Crucially “[v]ariation among these alternatives is not literally free; actually, since they differ in their autonomy, they also differ in the degree of freedom with which they are employed” (Lehmann 2002:310), which partially determines possible ensuing diachronic processes. Semantic weakening occurs in later stages of grammaticalization whereas earlier stages show “a redistribution or shift, not a loss, of meaning” (Hopper & Traugott 2003:94; also Bybee & Pagliuca 1987; Langacker 1990; Bybee *et al.* 1994).

The two important dimensions for this study are shown in Table 5.⁸ In (9), discourse factors generate variation between weak and strong pronouns, whilst pragmatic forces cause

⁷ Variouslly termed *scales* (Lehmann 1995), *channels* (Givón 1979), *chains* (Heine *et al.* 1991; Heine 1992, 2000), and (grammatical) *clines* (Hopper & Traugott 2003).

⁸ Discourse in is not accepted by everyone, here we follow Givón (1979).

movement of weak pronouns to second position, providing the setting for later reanalysis into modern clitics. In (10), Latin pronouns certainly weakened to become simple clitics, some authors believe that they went further, becoming morphemes (§1.1.2).

Table 5

9	discourse	syntax	morphology	morphophonemics	zero
					➔
10	lexeme	clitic	derivational affix	inflectional affix	zero

Whilst cognitive/pragmatic processes of metaphor, metonymy and context-induced reinterpretation are grammaticalization’s ‘means’, *reanalysis/analogy* are its driving force (Heine *et al.* 1991; Traugott & Heine 1991; Traugott & König 1991; Hopper & Traugott 2003; Bybee *et al.* 1994). They “do not define grammaticalization, nor are they coextensive with it, [but it]...does not occur without them” (Hopper & Traugott 2003:69).

Reanalysis indicates structural changes affecting an expression (or class of expressions) without significant surface-form alteration, occurring when hearers interpret an expression’s structure/meaning differently from the speaker (Langacker 1987:58). This requires that (at least) two possible interpretations/analyses are available. Reanalysis is covert, revealed only “ex post when the construction behaves in ways that presuppose its new structure” (Lehmann 2004:162). Reanalysis of hamburg+er ‘food from Hamburg’ as ham+burger became overt when forms such as cheese+burger become productive. This highlights the role of interpretation over absolute meanings of units, and the *need* for overlapping form/function pairs in language. Far from presenting difficulties *vagueness* (as opposed to ambiguity) is a *positive* property of language.

Analogy is “the attraction of extant forms to already existing constructions” and operates overtly, e.g. extension of suffix –hood (<had ‘person, condition, rank’) to contexts without human referents e.g. falsehood. Whilst reanalysis leads to linguistic innovation, analogy spreads innovation across systems: “reanalysis and analogy involve innovation along different axes. Reanalysis operates along the syntagmatic axis of linear constituent structure. Analogy, by contrast, operates along the paradigmatic axis of options at any one constituent node” (Hopper & Traugott 2003:63-64).

Lexicalization has received divergent interpretations (Brinton & Traugott 2005:ch.2) due to its close relationship with grammaticalization (e.g. Moreno Cabrera 1998; Lehmann 2002; Himmelmann 2004). Whilst grammaticalization tightens the internal relations between members of constructions, lexicalization makes them irregular and eventually eliminates them, by removing constituents from analytical processes: “[a] sign is lexicalized if it is withdrawn from analytical access and inventorized” (Lehmann 2002:1). Grammaticalization and lexicalization are orthogonal, which can “apply alternatively to a construction, but successively to an item” (Lehmann 2002:4). The crucial difference is that “[g]rammaticalization involves...analytic access to a unit...lexicalisation involves a holistic access to a unit, a renunciation of its internal analysis” (Lehmann 2002:13). Whilst clitics have clearly been grammaticalized, some analysts consider many combinations to have been lexicalized i.e. removed from analysis, and thereby inherently ‘arbitrary’.

1.3.2 Lexicalization of Italian Clitics

Italian presents a rich set of clitics with many putatively unanalysable usages, but has received little study under grammaticalization/lexicalization perspectives beyond limited references which go little further than its acknowledgement (Berretta 1985a, 1985b, 1989; Sala-Gallini

1996; Berruto 1985a, 1985b, 1986, 1987; Salvi 2001; Cennamo 1999, 2000; Nocentini 2003a, 2003b). Russi (2008) stands out for its lengthy study of such patterns, attempting to decompose grammaticalization into sub-processes culminating in lexicalization.

Russi (2008:7) considers that “these sub-processes pertain to specific clitics or clusters which completely lose their pronominal function and become fully incorporated into specific verbs. They thus involve both grammaticalization of the clitic pronoun into an obligatory morpheme and lexicalization of the verb-clitic constructions...into a single lexical unit.” Russi (2008:9) identifies “two main classes of clitics...anaphoric (pronominal) and discourse pragmatic vs. semantic-pragmatic/lexical or strictly grammatical”, distinguished primarily by the fact that standard analyses cannot explain “strictly grammatical” usages.⁹ Russi explicitly avoids consideration of nominative clitics (Benincà 1999; Poletto 1993, 1999; Benincà & Poletto 2005; Vanelli 1985; Rizzi 1986; Gorla 2004; *i.a.*) because they “do not seem to participate in semantic-pragmatic phenomena comparable to those observed for object clitics” (Rusi 2008:10) and simply ignores non-argumental datives. Such limited coverage brings into question the criteria for the distinction of two classes. Moreover, the argumentation points to inadequacies in “standard analyses”, rather than justifying the addition of further mechanisms to hide them.

The purposes for which ‘strictly grammatical’ clitics are employed, have been productive over centuries, but there is no evidence of the so-frequent-as-to-lead-to-lexicalization phrases which engendered them. This might be due to lack of source material. In more recent cases, however, it should be possible to observe their genesis. No such evidence is provided. Nor can modern cases be processes of analogy with older forms, since neither old nor new sets are

9 Chapter 4 for similar arguments concerning reflexives and their relationship to non-active voice.

sufficiently frequent. Moreover, the arrangements found in Italian are echoed in other Romance languages (e.g. Catalan, Espinal 2009; French, Abeillé *et al.* 1998). For these languages to arrive at such common positions (modulo availability of adverbial clitics) after a millennium of independent development, makes lexicalization an unlikely mechanism.

11	<table><tr><th>XVIII</th><th>XIV</th><th>XV</th><th>XVI</th><th>XVII</th><th>XVIII</th><th>XIX</th><th>XX</th></tr><tr><td>5</td><td>5</td><td>4</td><td>5</td><td>5</td><td>13</td><td>40</td><td>60</td></tr></table>	XVIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	5	5	4	5	5	13	40	60	
XVIII	XIV	XV	XVI	XVII	XVIII	XIX	XX											
5	5	4	5	5	13	40	60											
12	indovinar+la	XVI			indovinar+ci	XX												
13	correr+ci	XVIII			correr+ce+ne	XX												
14	contar+la	XIX			contar+le/se+la	XX												
15	dar+ci/la	XVI			dar+lo/sela/sele	XX												
16	andar+ne	XIII			volverne	XX												
17	rigirar+la	XVIII			rigar+sela	XX												
18	sbarcar+la	XIX			sbarcar+sela	XX												
19	menar+selo	XVI			menar+la	XX												

With general caveats as to the accuracy with which first attestations of ‘pronominal use’ can be determined, Viviani (2006) provides a history of initial attestation of such forms as found in GRADIT (11). As Viviani notes, there is no correlation between patterns shown by the same verb (12-15) or across verbs (16). Whilst attestation appears to generally follow complexity (17-18), this is not necessarily the case (19). The only definable trend is that recorded usage increases with time. All the patterns currently attested with at least one verb have been available since at least XVI^c. The greater the population using what is a relatively new language (initially spoken form and eventually written¹⁰), the greater the number of *recorded* uses of new V+CL patterns. Given the numbers for the last two centuries, it is clear that the phenomenon is highly productive. GRADIT also treats many uses as ‘obsolete’ i.e. common usage ebbs and flows with time. These patterns form a healthy ecosystem, not a moribund element of the grammar/lexicon.

10 What GRADIT presents as ‘Italian’ before the last century is largely the Tuscan literary language.

It is also necessary to take into account the conservative nature of dictionaries. Masini (2008)'s survey of the ItTenTen10 corpus for *-sene* cases discovered uses with many verbs not found in GRADIT. Viviani (2006) further illustrates that not only does the number of patterns vary between dictionaries, but also the accepted uses of those patterns i.e. attestation is a biased choice on the part of lexicographers. In reality, many of these usages may have occurred for a long time and simply not been recorded as such. This all argues against a process of progressive grammaticalization, even less one of lexicalization/fossilization.

Currently the combined meaning of *-sene* allows it to be added to all motion verbs. Unless new roots are introduced into the language (very rare in this set), new coinages are impossible. The lack of such new forms does not imply anything about the mechanism's productivity: the class to which it applies is complete. Moreover as shown in §5.5.6, *-sene* is applied to 'unusual' verbs (with the same compositional meaning) in specific one-off situations. These are not documented in dictionaries because they never become sufficiently frequent or widespread, but are discoverable in modern corpora. The existence of such cases indicates that composition is productive, as far as is possible to its meaning. The reason that similar patterns develop across Romance languages is, we believe, due to the similarity of meaning in the individual clitics and a common process of composition.

Masini (2008) presents a wide range of uses, where the lexicalized group *-sene* as a unit within a *Construction* pattern can be applied to new verbs if the new usage overlaps sufficiently with existing stored uses. However, there is little advantage to such an approach if direct composition remains available. In order to become stored as lexical entries (*à la* Russi) or constructions (*à la* Masini), units must be frequent. speakers must, therefore, have been

able at some point in history to regularly compose these forms. It behoves lexicalists and constructionalists to (1) explain their compositional meaning at that earlier stage and (2) explain how/why/when this meaning~form pair became so opaque as to require lexicalization, as lexemes or constructions. Neither element of argumentation is addressed by these authors, or any other which we could find.

Chapter 5 sets out to show that the cases presented by Russi and Masini are better explained compositionally, by extending the analysis of what functions clitics may perform i.e. dealing with those inadequacies in “standard analyses” and without the need to add intermediary mechanisms such as lexicalized *Constructions* or lexical entries. It provides a compositional analysis of *-sene* and other ‘difficult’ combinations, finding no evidence for any change in the transparency of their composed meanings. Rather, it is only by keeping each of its elements as separately applicable, that it is possible to understand the full range of uses of *se*, *ne* and *sene*.

This work starts from a position which rejects the removal of any clitic (or combination thereof) from the analytical process and its lexical storage as being unnecessary, and hence an added burden/inefficiency which biological systems tend to eschew. We seek explanations purely in terms of the functions which a clitic may perform and the composition of those functions with each other and the verbal context; until it is proven that something more is necessary. It may be that, in a wider concept of language, further layers of abstraction are required, as proposed by constructionalists. In the case of clitics, we find no evidence for their necessity nor usage. For the purposes of this work, therefore, we take the strong position that such composition takes place purely at the semantic level and is expressed through structure which we can recognise from surface form i.e. there is a direct link between the message and

its content which can be learnt purely by positive experience and extended by speakers to new environments where appropriate.

1.4 A Communicative Approach

In Cognitive Linguistics (e.g. Fillmore 1985; Lakoff 1987; Langacker 1987; Talmy 2000), language is not autonomous from cognition, rather its structure is explained by reference to cognitive principles and mechanisms, such as general categorisation, pragmatic and interactional principles, which underpin human conceptualisation of the world, not just language. Language is how humans construe reality (Haiman 1980, 1983). Three basic principles of compositionality, inference, and iconicity combine to explain variable acceptability, negative exceptions and acquisition.

The speaker's task is to project non-linguistic experiences onto linguistic expressions, matching his construal of experiences to conventional values of linguistic symbols, chunking the experience into a small number of "things talked about" (Gentner 1983, 1988). Such experiential chunks and their inter-relationships are structurally mapped (Gentner & Markman 1997; Gentner *et al.* 2001; Fisher 2000; Kako 2006) in order to recognize "things talked about" in their proper inter-relation. To retrieve the speaker's message, the hearer must perform reverse cognitive mapping. Since language users act as speaker and hearer, they benefit in one mapping from their knowledge of the other (Hurford 2003; Hawkins 2004:25). Indeed, self-corrections suggest that linguistic production involves analysis by synthesis, matching mapping of articulation with envisaged hearer de-mapping (Keller 1995:180-181).

Repeated use of the same chunking results in common linguistic symbols i.e. stably shared recurring partials (Tomasello 2003:51). Since these symbols are language-specific categories

abducted by general cognitive skills of pattern finding (Tomasello 2005:191-194; Bowerman & Choi 2003:407-409), it is unnecessary that “the structure and principles of CS [conceptual structure] are present in the learner prior to the task of language acquisition” (Culicover & Nowak 2003:11). Furthermore, practice in specific (re)chunking, will eventually come to guide “chunking” of experience (Loucks & Baldwin 2006, 253). Similarly Lucy 1992:275; Lucy & Gaskins 2003; Gentner & Goldin-Meadow 2003:10-11; Gentner 2003:225-28.

Given that speakers cannot provide more than weak outlines of their construal of a situation, hearers are required to integrate new information evoked by the speaker’s sparse hints with their own background knowledge (Sperber & Wilson 1986:153; Bransford & Franks 1972:221-5; Sanford 1999:304; Garrod & Pickering 1999:3), and arrive at contextually coherent conclusions (Elman *et al.* 2005:111); words are merely “abstract constraints that guide meaning-making acts” (Bransford & McCarrell 1977:396). As Wright (1976:519) observes, “there is no guarantee other than the ‘utterer’s’ and ‘hearer’s’ common satisfaction over their mutual pragmatic success that they are taking their meanings in the same way.”

Communication is made possible by human problem-solving capabilities, combining clues and drawing conclusions (Levinson 2000). It follows that symbol-combinations are *only* interpretable in context (Deacon 2003:129-33) which is confirmed by experimental evidence where hearer re-construction is facilitated when context confirms his inferences (Murray & Liversedge 1994:366-68; Tyler & Marslen-Wilson 1977:684-5; Tanenhaus & Trueswell 1995:239-41; Boland 1997:609-10; Britt *et al.* 1992:302; MacDonald *et al.* 1994:678).

Speakers are facilitated by a close relationship between the symbolic sequence and the experience being communicated. Hearers are facilitated, the more iconic the expression (Bock

1982:6,13,35; Fisher 2000:19-20; Newmeyer 2001:104; Deacon 2003:124). Minimising the cost of “processing enrichment” (Hawkins 2004:44-48) is key to easy communication (Newmeyer 2005:1669). Hence messages characteristically display motivational or “diagrammatic” iconicity (Kleiber 1993:106; Haiman 1985:9; Hollmann 2005:288-90): “we keep finding iconicity because there is no other way for a semiotic system to be created and used by human beings without a close fit between form and function” (Slobin 2005:320).

1.4.1 Explaining Exceptions

Corpus analysis and experimental work show that relative frequency of syntactic alternatives varies along semantic, syntactic, lexical, and phonological continua (Wasow 1997; Wasow & Arnold 2003; Gries 2003; Arnold *et al.* 2004; Lohse *et al.* 2004; Gries & Stefanowitsch 2004). Thus, choice between English genitive *'s* vs. *of* depends on the relation between the two entities, and hence factors such as relative topicality, animacy, concreteness (Deane 1987; Rosenbach 2003; Stefanowitsch 2003). Manning (2003:319-22) suggests that (in)frequency continua culminating in the absolute non-occurrence of variants can be formally modelled within probabilistic syntactic frameworks without substantive motivation, however, such ‘distributional constraints’ merely label non-occurrence, rather than explain it (Jurafsky 2003: 93-94). “Frequency effects as such do not constitute an explanation but are themselves an effect of more general and processing-related principles” (Verstraete 2005:501).

Syntactic processing does not provide such clear motivation as that found in phonetics, where articulatory/perceptual considerations facilitate explanation (Browman & Goldstein 1992; Lindblom *et al.* 1995; Lindblom 1999; Pierrehumbert 1999:295; Broe & Pierrehumbert 2000:7). Nevertheless, it is possible to consider unequal cognitive costs of formulations.

Cross-linguistically, structures which are easier to process are more frequent (Kirby 1998:365-66); familiar and/or prototypical items receive higher grammaticality judgements (Manning 2003:301-2; Bybee & Eddington 2006; Scholz & Pullum 2007:715; Stefanowitsch 2008:527); whilst in syntactic variants which differ in length and, therefore, amount of real-time processing, the cognitively more economic alternative is favoured (Hawkins 2004).

Expressive alternatives may be explained in similar fashion. Referentially equivalent variants require different computations with unequal cognitive costs (MacLaury 1991; Stubbs 1996:215). Since alternatives present the scene from different perspectives, each will be unequally congruent with different contexts (Maiden 2004:253). This approach provides a coherent argument not only for why given patterns generally fail to occur, but also why in exceptional contexts, and for very infrequent communicative needs, proscribed combinations do occur (Stefanowitsch 2007:68). As García (2009:15) illustrates, (20) is normally considered an unacceptable version of (21). Nonetheless, (20) proves acceptable in (22, *Egoist* p.489), and more appropriate than (23), because it occurs within Meredith's work as a whole, which manifests Sir Willoughby's morbid dependence on images others have of him.

Table 6

20	**Himself killed him
21	He killed himself
22	"of Sir Willoughby; he was thrice himself when danger menaced, himself inspired him."
23	He inspired himself

Systematic avoidance reflects arrangements so difficult to interpret and/or requiring so much contextual support, that an alternative form better serves communicative needs, at lower cognitive cost to speaker and hearer (Newmeyer 2005:1669). "It is not that the English language (or any other language) presents us with a fixed finite range of constructions which

rigidly constrains our linguistic behaviour; rather, our speech and writing make heavy use of the best-known patterns of the language, but we are free to adapt these and go beyond them as we find it useful to do so, and there are no such things as word sequences which are absolutely “ill formed in English” – only sequences for which it is relatively difficult to think of a use, or for which no one happens yet to have created a use” (Sampson 2001:166).

1.4.2 Non-Arbitrary Spanish Clitic-Clusters

In a detailed study of Spanish clitic-cluster anomalies, García (2009:2-3) argues that “what matters is the communicative value of individual signs, and the mental calculus required to interpret symbol combinations...the acceptability of a clitic combination depends on whether the cluster is interpretable in the sense suggested by its context, given the constraints imposed by real-time processing...this allows a principled account of the notorious rejected clusters.” For García (2009:291), it is “difficult to reconcile contradictory or incompatible inferential manoeuvres” in certain combinations: “time-consuming computation” leads to their rejection.

García successfully shows that such motivation does exist for many Spanish anomalies. Throughout this work, we provide examples where (a) certain usages are less frequent, some to the point of (almost) never occurring, but can do so given appropriate context; (b) genuinely ambiguous clusters (due to shared surface-forms) are generally avoided, along with cases where they are used in real life leading to requests for clarification, proving that they are not impossible, merely communicationally ineffective (e.g. §3.5.2); (c) complex clusters which include normally avoided combinations, exactly because those combinations cease to be ambiguous in those contexts. The approach also encompasses normative prescriptions e.g. **me+se* (§1.2.3), where avoidance is based upon what usage says about oneself. This is

simply another kind of evaluation of a signal's communicative worth. For negative exceptions, it is inappropriate to consider syntactic or morphological constraints of the *X+Y type; clusters are used when they are meaningful and not when they fail to communicate. Users of a language know when this will occur and choose the most effective variant.

Less convincingly, García attempts to explain positive exceptions by combination/interaction of the arguments used to explain negative ones. The central issue with 'cognitive economy' is that, just as with surface templates, analysts know the desired results and so create rules to achieve them, rather than observing patterns emerging from independently motivated models. García's argumentation is based upon unfounded presumptions. Difficulties are not necessarily cumulative and even if they were, it is unlikely that they are equally weighted i.e. it is not possible to simply add them up and stop using forms above certain difficulty count. There is no evidence of 'computation bottleneck', just as the desire to reduce forms to minimise feature count for reasons of space has no basis in memory limitations. The implication is that the whole message is being transmitted along an insufficient pipeline. However, it is the nature of speech that it does not attempt to express everything, merely provide hints for re-creation. As a compression technique, clitics act as references back to in-stream data already analysed by both speaker and hearer. Such zipfian compression requires minimal processing; indeed that is its *raison d'être*. With respect to memory, the state variables requiring storage are minimal: 1/2-persons are defined by the conversation and always available, only 3-person is in question, i.e. how many 3-persons can be maintained and to what depth.¹¹

11 Helping to explain why so many issues revolve around 3+3-clusters.

García (2009:37) offers impersonal *se*'s inability to co-occur with another *se* as an example of extending the analysis to positive exceptions. However, there are perfectly reasonable structural arguments which not only explain this, but also why the Italian equivalent is allowed but mutates to *ci+si*, and dialect *si+si* (§4.6.9). García's approach, by definition, cannot deal with these cases, since the analysis will always disallow such cases due to cognitive cost. It has been argued that inferential routines are rooted in language-specific evaluations (Dryer 1997; Levinson 2001; Fortescue 2002; Everett 2005). Grammatical meanings emerge as obligatory contrastive categories from the frequency with which a particular categorisation is made (García & van Putte 1987), and thus, any universality in content reflects the similarity of communicative needs across human communities, just as formal universals reflect semiotic constraints (Deacon 2003:126-34). In this case, Spanish and Italian world-views, and the nature of the languages which they have engendered, are too close to presume wholly different inferential rules, whilst the same Italian speaker may use *ci+si* or *si+si* depending on social context i.e. whether national or local dialect is most appropriate. García's equations of relative interpretational difficulty and their combination as an explanation of positive exceptions must, therefore, be questioned.

There are also qualitative differences between negative exceptions which may be reversed with adequate contextual support, and positive exceptions which cannot. In these cases, clitics are not at extremes of any plausible continua, some are acceptable or unacceptable when expressed as full arguments (i.e. they do or don't represent variations), and there is often no ambiguity to resolve, so there should be no problem of interpretation.

That negative exceptions can be explained as cognitively motivated, does not rule out other factors being involved. It simply means that analysts have to distinguish more carefully between those cases which are truly motivated (negative exceptions) and those where motivation is indirect (positive exceptions). Many PCC exclusions break the logic of interpretability motivating syntax, unless we enter the world of somewhat forced cumulative evaluations which do not hold cross-linguistically. There must to be another dimension which ‘prevents’ these occurring, and forces the speaker to alternative formulations (even though this limits choice of expression).

This work considers that limiting factor to be syntactic structure, about which García avoids discussion. This does not deny the relevance of García’s arguments, but rather abstracts them to a higher level. Semiotic systems are iconic by their nature, and any syntactic structure developed to express that system will naturally reflect this, but being subject to other constraints, only indirectly. The same general motivation is the source of all restrictions, some directly at the level of cognitive analysis where clitic referents are obtained and some at the level of syntax, through which mapping form to and from function occurs.

1.5 Conclusions

From our perspective, all the approaches discussed above share a premature acceptance of arbitrariness: either as a formal statement of intent (*autosyn*) leading to approaches where clitics are seen as an “irrelevance to pure formal syntax” (Wanner 1994:51) and attributed to “weird morphological constraints” (Bonet 1994:51), or implicitly by virtue of extraction to “unanalysable chunks”, treating the lexicon as “a trove of irregularity” (Newmeyer 1998:219).

This work starts from the premise that all clitic behaviour has a sound reason, until proved otherwise i.e. we believe that each clitic has a definable function, reflected in syntactic usage, whereby the meaning of a phrase is merely the composition of those functions and the verbal context.

Essentially, this work is attempting to define the target description against which any detailed syntactic model can be measured for empirical adequacy, rather than the processes by which each clitic arrives in its position. It is, therefore, irrelevant whether they are base-generated or products of movement. Such details are not important to the *what* and *where*, but only the *how*, of clitic positioning.¹² We won't be proposing complex theories. Exactly the opposite. We argue that simple structure allows us to meet the full (not idealised) data, without the need for most *X+Y style exclusions, X+Y→Z+Y conversion rules, complex interpretational mechanics, or unanalyzable (and, therefore, lexicalized) units. By focusing upon the three principles of compositionality, inference, and iconicity (introduced above and developed in successive chapters), it is possible to define a system which is learnable whilst returning clitics to their rightful place within the heart of syntax.

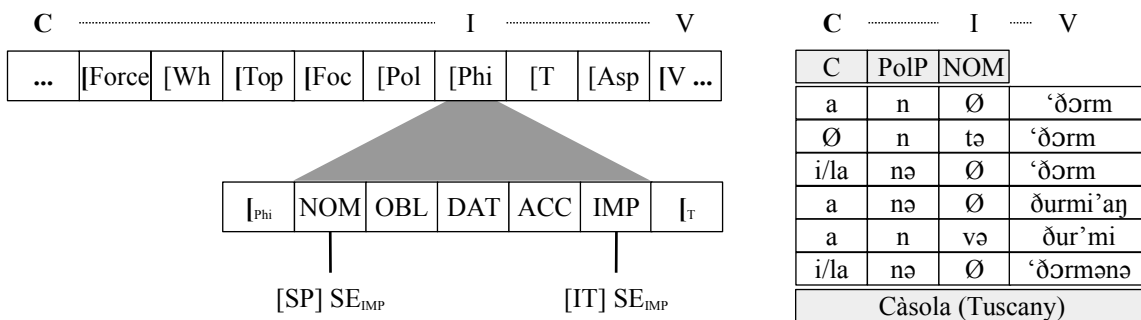
¹² Equally, whilst we explore several historical sequences of change, lack of space precludes investigation of extra-linguistic forces which may have influenced such changes.

2 MODEL

A key tenet of the current proposal is that clitic-forms may perform more than one syntactic function, reflected in their position within an elaborated series of feature projections including heads, not only for VP argument referents, but also non-argumental datives and nominative actors. Surface clitic patterns are merely sequential spell-outs of this structure. Giving each case its own position reduces the need for exclusions and inter-clitic processes, leading to a simple and coherent model readily applicable across Romance.

2.1 Elaboration of ‘Standard’ Models

In the C-domain, sentence grammar meets discourse (Rizzi 1997; Benincà & Poletto 2004). C’s left-edge encodes sentential ‘force’ (declarative/interrogative/exclamative) attracting *wh*-phrases and exclamative elements,¹³ followed by topics and point-of-view constituents, such as discourse-linked (‘contrastive’) focus. C’s right-edge (Pol) asserts/denies propositions.



The I-domain hosts verbal inflectional constituents (tense/aspect) with Phi projections at its left-edge, immediately below Pol. In Old Romance (e.g. Old Spanish, Rivero 1991) clitics or

¹³ Further topics above ForceP occur in root clauses e.g. Spanish *¿a María_{TOPIC}, quién la invitó?* ‘Maria, who invited her?’

possibly WPs (weak pronouns) could appear in C- and/or I-domain, but are restricted to I-domain in most modern languages. Some Gallo-/Italo-Romance dialects retain some topic/focus C-clitics, separated from Phi clitics by Pol.¹⁴

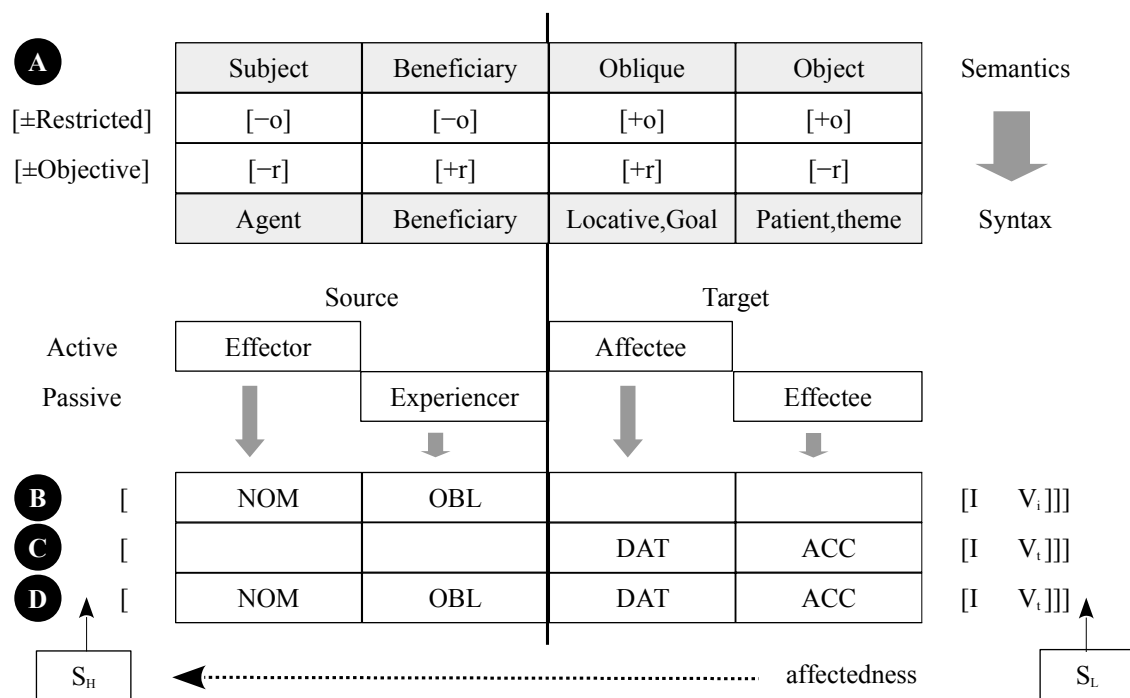
This model subdivides Phi. Whilst ACC (accusative) and DAT (dative) roughly correspond to direct- and indirect-objects, NOM (nominative) and OBL (oblique) host non-VP arguments. The traditional term ‘ethical dative’ is inadequate, since it masks distinctions between *dativus (in)commodi* vs. *ethicus* and between event affectees (OBL) vs. effectors (NOM). §4.7.1 differentiates the latter based upon semantics, syntactic behaviour and relative position.

One further position is required. Italian SE_{IMP} (used to identify indefinite subjects) appears between ACC and V. Whilst Italian had developed SE_{IMP} from SE_{PASS} before the earliest texts, languages which developed SE_{IMP} later (e.g. Spanish) grammaticalized different usages such that it now appears under NOM. §4.6.9 explores these and further variations and their effects on cluster availability.

As illustrated in (A), the proposed projections match functional classifications of Lexical Mapping Theory (LMT, cf. Bresnan & Kanerva 1989; Bresnan & Moshi 1990) which postulates two features, constraining the mapping of semantic roles onto grammatical functions. In a case-model, the dividing line exposed by [\pm o] also reflects structural division.¹⁵

¹⁴ This often referred to as NegP. Here, NegP is treated as the realization of a more general polarity phrase, PolP (Laka 1990) hosting elements that negate (e.g. Spanish *no/nunca*) or assert (e.g. Spanish *sí/siempre*) propositions.

¹⁵ Similarly, Comrie (1981:53-6)’s control continuum places experiencers closer to agents, and separated from patients.



IP is seen as forming two distinct fields, each containing two participants in an asymmetric relationship where the dominant partner is actively involved in the construal and the subordinate is an experiencer at that level. Whilst intransitives support only the upper field (B), transitives also license transitive sub-structures (C). These fields also stand in an asymmetric relationship (D) where source (impetus into the event) dominates target (external entity acted upon).

The cumulative effect of these relationships is that the sequence of elements within the verbal frame is an iconic representation of participant ‘affectedness’ within the construal. The verb acts directly upon the least active participant (ACC), indirectly affecting its dominant partner (DAT) e.g. due to loss/gain of possession, whilst the action itself (i.e. transitive sub-structure) affects its dominant partner (OBL), e.g. a (di)transitive action, of which (s)he is not a part, is carried out for his/her benefit. Finally, effectors (NOM) may be affected by the process which

they have set into effect, often seen in terms of satisfaction (§4.7.2). The participants, therefore, represent a chain of decreasingly direct affectedness, reflecting the empathy scale (Givón 1984).

Subjects may appear in two positions: S_L (low) which may be associated with the initial merge site of the verb's external argument and is seen in so-called 'subject inversion' constructions; S_H (high), the canonical position for subjects in declarative sentences, usually associated with movement to SPEC,IP or higher in the C-domain. The availability of two positions 'continues' the scale. S_L is generally reserved for inactive subjects, whilst raising to S_H requires agentivity. Note that the notion of agent used here, profiles ability to perform actions by virtue of inherent properties; Higginbotham (1997)'s "teleological capabilities." Some inanimates, or non-intentional animates, may be agents in this sense ("theme unergatives", Levin & Rappaport 1995) e.g. verbs of sound emission, *The train whistled*. Similarly, unergative verbs like *cough/blush*, whose subjects are animate, but rarely intentional.

There have been numerous proposals along similar lines. For example, Sportiche (1995) considers clitics to be generated in functional heads within tense corresponding to AgrS (\approx IP,SPEC), AgrIO (\approx DAT), AgrO (\approx ACC), immediately above ν P. This case-model includes OBL, and NOM clitics as first-class members of the set. Manzini & Savoia (2004) argues that clitic heads form subject~object clusters repeated above C/I/V. Each head may host D-features, ϕ -features, and possibly case-features, which may be lexicalized by full DPs or clitics. In a case-model, clitics (sets of ϕ -features) are *hosted* by case-ordered heads, where case (NOM/OBL/DAT/ACC) is defined by participant~participant and participant~event

relationships, whilst pairs (NOM~OBL and DAT~ACC) are defined in terms of direct~indirect rather than subject~object relations.

Not only is the elaboration envisaged by a four case-model small, but it brings with it connections to semantic (LMT) and cognitive (empathy scale) models which would otherwise remain disjoint, whilst defining case in terms of the structural relationships of which these are the surface realization. It is our contention that separating NOM/OBL from DAT/ACC is central to a working model of clitics. It is only by accepting the presence of NOM/OBL as equal partners that we can clarify the range of combinations/processes in DAT/ACC and ultimately provide an adequate explanation of them.

2.1.1 The Current Model

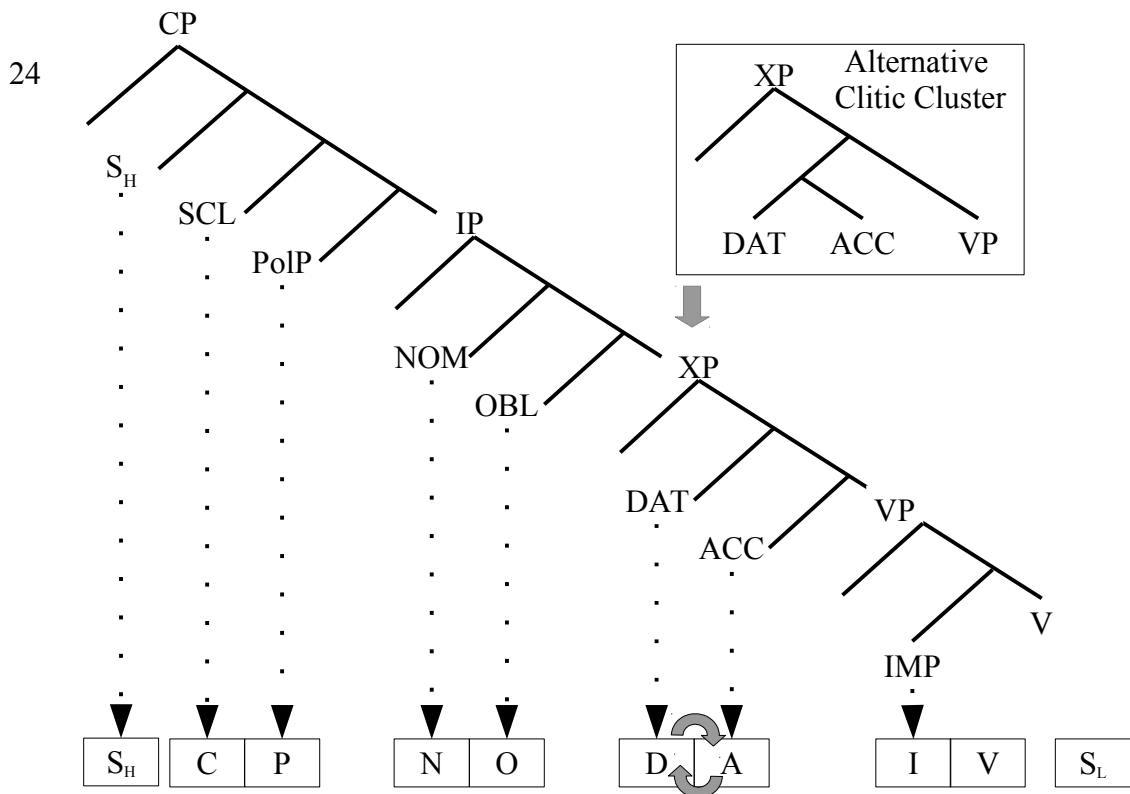
The basic pattern is presented in (24). CP/IP/VP are convenient labels without implying support for, or reliance upon, any particular theory; indeed our use may conflict with some proposals, e.g. Zanuttini (1997) considers PolP (referred to as NegP, see fn.14, p.33) to be on IP's left-edge, rather than CP's right-edge. The essential point is that the element sequence is syntactically fixed, divided into two sections, and reflected directly in surface sequences.

CP	hosts	S _H (optional in pro-drop languages)	relating to the	clause
IP	hosts	NOM/OBL (non-object arguments)	relating to the	event
XP	hosts	DAT/ACC (object arguments)	relating to the	action

The presentation is schematic, excluding material (e.g. adverbs) irrelevant to the discussion. The detailed shape of each block (CP/IP/XP) is unimportant; each pair might form an applicative structure (Pylkkänen 2002) rather than the shells illustrated, or use alternative cluster formations (Ordóñez 2002). The central issue is recognition of NOM, and its pairing

with OBL. There are several high applicative analyses of Romance and Slavic languages,¹⁶ however, all treat NOM clitics as equivalent to ‘ethical’ datives. We argue that a comprehensive and coherent model requires a four-case design.

In our model, the *only* surface sequence variation is D/A-swapping (indicated by the curved arrows), as historical and synchronic processes.¹⁷ Chapter 6 shows that swapping is based on individual clitic ‘weights’ e.g. in French, heavy *lui*⁺_{DAT} causes *lui*⁺_{DAT}+*le*_{ACC}→*le*+*lui*⁺. There is no evidence for N/O swapping. Most NOM clitics are ‘light’, whilst OBL has some ‘heavy’ clitics. It may be that no combinations warrant alternation, or that this difference between D/A and N/O is indicative of structural differences which future models should reflect.



16 e.g. Roberge & Troberg (2009, French); Bosse & Bruening (2011, French); Cuervo (2003, 2010, Spanish); Grashchenkov & Markman (2008, Russian); Folli & Harley (2006, Italian); Diaconescu (2004, Romanian).

17 e.g. Old French’s rigid A/D ordering has developed to D/A order in some regional dialects, whilst standard modern French shows mixed properties (Morin 1979).

Within the cartographic tradition, Poletto (2000) models the subject clitics of Northern Italian dialects on the premise that 1/2-person clitics occupy a distinct position from 3-person. Similarly, Bianchi (2006) for Italian object-clitics. Equally, it has been argued from differences in c-command relationships in French (Boneh & Nash 2011) and Spanish (Cuervo 2003) that lower benefactives are syntactically higher than goal/recipient arguments, although still within VP as indicated by interaction with the PCC. Nevertheless, 1/2-clitics are mutually exclusive with 3-person clitics, as are lower benefactives with all other dative/locative uses. For the task at hand, it is sufficient to work on the basis that such mutual exclusivity indicates single syntactic positions, where further distinctions such as reflexivity are treated as features of that position, i.e. semantic subtleties may be reflected in a node's exact position, but each node and thereby its clitic remains the 'representative' of its block, and therefore (modulo A/D-swapping) in a constant sequence. Thus, we repudiate the central tenet of Perlmutter and much subsequent work, that there is no underlying structure which can explain surface forms. In case-models, semantic functions are reflected transparently in syntactic structure, and surface form is merely its sequential spell-out.

2.1.2 Items Not Considered

SCLs (a term used somewhat loosely in the literature to cover both C and N clitics, or combinations thereof) are common in northern Italy (Poletto 2000), Franco-Provençal (Roberts 1991), and Rumantsch (Linder 1987). In other languages, C clitics are Ø, with N alternating between Ø and reflexive forms. Space precludes discussion of the wide range of variations found across Romance of SCLs (Table 7 gives a few examples from Manzini & Savoia 2005). With the possible exception of 3-3-contexts, the literature does not discuss any relevant form changes, nor movement between N and O/D/A. The main research questions

revolve around division of such clitics between C and N and when they appear or surface as Ø, which varies across these languages according to various discourse properties. In this work, the N clitics of these dialects are treated simply as non-reflexive NOM clitics which happen to have developed a surface-form.

Table 7

C	P	N	C	P	N	C	P	N	C	P	N
(e) n(o) i 'dərma			i nun Ø 'dərme			a n Ø 'ðorm			a n Ø 'drəm		'næinta
(e) non tu 'dərma			tu n Ø 'dərme			Ø n tə 'ðorm			a n t 'drəmi		'mai
(e) no llə 'dərma			i/ε nun Ø 'dərme			i/la nə Ø 'ðorm			u n Ø 'drəm		'næinta
(e) non Ø dor'mjaŋ			Ø nun sə 'dərme			a nə Ø ðurmi'aŋ			a n Ø dru'muma		'næinta
(e) non Ø dur'middə			Ø nun Ø dur'mitə			a n və ður'mi			i n Ø 'drəmi		'næinta
(e) no llə 'dərmaŋ			i nun Ø 'dərmaŋ			i/la nə Ø 'ðormənə			i n Ø 'drəmu		'næinta
Sillano (Tuscany)			Vagli di Sopra (Tuscany)			Càsola (Tuscany)			Oviglio (Piedmont)		

25	Dze medzo-dzò an pomma?	Shall I eat an apple?	Valdôtain, Franco-Provençal (Roberts 1991:307)
26	i durmin	We are sleeping	Forni di Sotto, Friulian (Manzini & Savoia 2005)
27	durmin=os?	Are we sleeping?	
28	Ou migi sa soupe	He ate his soup	Limousin, Occitan (Doussinet 1971:391)
29	Migi-t-eu sa soupe?	Did he eat his soup?	
30	Ou(s) migeant	They are eating	
31	Migeant-î?	Are they eating?	

32		1.SG	2.SG	3.SG.M	3.SG.F	1.PL	2.PL	3.PL.M	3.PL.F
	CL+V	i	te	i/l'	le/l'	ne	os	i/l'	le/l'
	V+CL	ou	t'	u	le	n'/nous	ous	u	lè

Pontarlier, French (Tissot 1865[1970])

One area which will have import for future developments of the current model is the 'subject inversion' properties of these dialects, which cannot be simple cases of movement, since pre- and post-verbal SCLs may co-occur in some varieties e.g. Valdôtain (25). Cardinaletti & Repetti (2008) argue that one form is derived from the other through morpho-phonological processes, but some cases seem to require a suppletive analysis (26-27). In dialects of Limousin (Occitan), 3.SG/PL subject clitics are realized as pre-verbal *ou(s)* vs. post-verbal *t-eu/t-i* (28-31). Whilst [t] may be derived from an old liaison consonant (cf. French *-t-il*, Foulet 1921:269), synchronically, *t-eu/t-i* act as distinct post-verbal subject-marker forms. Similarly,

some Franco-Provençal dialects show pre-verbal *i(l)/l* vs. post-verbal *tì* (Olszyna-Marzys (1964:36). Finally, the French of Pontarlier (Eastern France), shows no systematic correspondences between pre- and post-verbal forms (201, Tissot (1865[1970])). It appears that some languages have a post-verbal clitic position for nominative clitics (as well as those languages which leave object-clitics after the verb). We tentatively assume that this position is related to the I position described for Italian.

Most Romance languages have single high or low (post-verbal) adverbial negators, or combinations thereof. Languages using lower adverbials may show further possibilities e.g. Càrcare (1), where ‘negative clitic’ *η* may be applied in various combinations. Zanuttini (1997) proposes four positions for such negative adverbs, where PolP is simply the highest and most commonly used.¹⁸ Since they do not affect our argument, i.e. they interleave with the proposed projections, such negators are not discussed further.

Table 8

1	C	↓	[N	↓	[D	↓	A	↓	[V		
	ε	η	tε	η	tε	η			[‘lɔvi nɛ:nt	You don’t wash yourself	(Càrcare, Liguria)
	ε		tε	η	mε	η			[‘ʃɔmi ‘mɔi	You never call me	(Manzini & Savoia 2005)
	ε	η	tε		mε	η	lε		[‘dɔi ‘mɔi	You never give it to me	
	u	η			sε				[‘lɔva nɛ:nt	He doesn’t wash himself	
	u	η			mε	η			[dɔ ‘nɛ:nt	He gives me nothing	
	u	η			mε	η	lε	η	[‘dɔ	He doesn’t give it to me	

We maintain traditional distinctions between *dativus (in)commodi* (2) and *dativus ethicus* (3). Whilst neither is sub-categorized by the verb, the *dativus ethicus* is limited to 1/2-persons, and not related to the event but the speech-act, designating persons taken as witness among the interlocutors. Woodcock (1959) translates 1-person *dativus ethicus* as ‘pray’ reflecting their

¹⁸ Negation and clitics are sometimes reported as ‘swapping’ e.g. Cairese (Ligurian, Zanuttini 1997). We take these reports to be cases of multiple negation positions rather than movement processes.

non-referential status. As Cardinaletti & Starke (1994:51) assert, *dativus ethicus* are discourse particles, and as such, “there is no referent to these pronouns, not even derivatively.” Joutiteau & Řezáč (2007, French), Salvi (2001, Italian) and Diaconescu (2004, Romanian) provide evidence that the two types are semantically and syntactically distinct. The literature, however, often ignores the differences, using various terminology: ethical/affected/non-lexical/dative-of-interest.

Table 9

2	Sol omnibus lucet	The sun shines for everybody (Petronius, Satyricon, 100, in Van Hoeske 1996:7)
3	Quid mihi Celsus agit?	How, pray, doth Celsus fare? (Horace, Epistulae 1,3,15, in Woodcock 1959:47)
4	Il te lui a donné une de ces gifles!	
5	Au Mont St Michel, la mer te vous monte à une de ces vitesses!	
6	Ce pleurnicheur, il te se met en larmes pour un rien.	

French ‘ethical’ datives are considered characteristic of ‘low registers’ but prevalent in some southern regional varieties (Charaud 2000:648). They may co-occur with other non-thematic datives (4), often in pairs (5, Leclère 1976:93), and sometimes transgress combinatorial constraints (6, Jones 1996:301). Their mobility is explained if it is assumed that they are truly adverbials able to take various positions (positive equivalents of the variable position negatives). Although exemplified at various points for contrastive purposes, this work does not consider them further. The OBL position of the current model hosts *dativus (in)commodi*.

2.1.3 Spell-Out

Each node is represented in surface-form in syntactic order (modulo D/A-swapping). Within the syntax-tree, clitics are defined for reflexivity [$\pm R$], and [$\pm E$]. The remaining features are derived from the referent (Table 10).¹⁹ Since each pair of [$\pm R$]/[$\pm E$] form mutually exclusive

¹⁹ Many of the table entries are filled in other languages (§2.2.3).

sets, they are treated in this work as featural differences, however, the ‘feature tree’ could also be expressed in more detailed syntactic structure, without significant changes to the approach.

Previous analyses tend to associate all uses which take dative forms. The current model not only makes a clear distinction between source- and target-domain ‘datives’ (OBL~DAT), but establishes two distinct functions for each ([±E]). For DAT, the distinction is between *affected* participants (traditional datives) and distal functions (spatial designations).²⁰ For OBL, the division reflects what are sometimes termed ‘sympathetic’ vs. ‘setting’ datives. [±E]’s definition is filled out in subsequent chapters: non-reflexives (Chapter 3), reflexives (Chapter 4), adverbials (Chapter 5). At this point, it merely represents the need for two categories as shown by the fact that some clitics are available under one heading but not the other.

Table 10

Spanish																											
Syntax	NOM				OBL				DAT				ACC				I	V									
	+R		-R		+R		-R		+R		-R		+R		-R												
	-E	+E	-E	+E	-E	+E	-E	+E	-E	+E	-E	+E	-E	+E	-E	+E											
Referent	me								me				me				me				1				Singular		
	te								te				te				te				2						
	se								le				le				lo				3M						
									Ø				la				3F										
													lo				3N										
	nos								nos				nos				nos				1				Plural		
	os								os				os				os				2						
	se								les				les				los				3M						
									Ø				las				3F										
																los				3N							
se				se				Ø				Ø				Ø								[SPEC]			
N				O				D				A				I											

20 On dative~locative proximity, Jespersen (1924:ch.XIII).

As the contrasts between Italian NOM_[+E] *ci~vi~Ø* (§5.4.6) and DAT_[+E] *ci~vi~glie* (§6.2.6) show, [–SPEC] clitics also show person (i.e. proximal~medial~distal) distinctions.

Many analyses invoke a common understanding of reflexivity: e.g. Seco (1988:199): “*la acción verbal vuelve como un rayo de luz en su espejo sobre el origen de donde procedió*”. Since many uses do not seem to fit, some reject this basic metaphor. RAE (1973:§2.5.5) repudiates *reflexive* as semantically equivalent to ‘actions directed to oneself’, treating it merely as grammatical concordance i.e. subject co-reference. However, co-reference is clearly inadequate as this would subsume all subject pronouns, leaving no distinction between *a él/sí mismo*, and no means to express coreferent, but non-reflexive, clitics as seen in Northern Italian dialects, or Spanish impersonal *se*, which are clearly coreferent, but by no means reflexive.

Clitics may be coreferent, with/out being ‘reflexive’. The relationship is shown in (7). Whilst all NOM clitics are subject coreferent by definition, only those marked [+R] are reflexive, i.e. require ‘reflexive’ forms, paralleling the contrast between subject pronouns, *Yo~Yo mismo* ‘I~I myself’. This is surface evident in Northern Italian dialects such as Càrcare (8, Liguria, Manzini & Savoia 2005), where NOM_[–R] (which is subject coreferent not reflexive, as shown by 9) coexists with DAT_[+R], even though both are subject coreferent.

Note that since OBL is not a verbal argument,²¹ it cannot be subject coreferent, nor less reflexive. §7.5.5 shows that this property emerges from structure.

21 Perlmutter (1971) and Jaeggli (1982) for numerous arguments and examples.

Table 11

7 [REFERENTIAL [COREFERENT [REFLEXIVE]]]

	C		N		D		A		
8	ε	η	tε _[-R]	η	tε _[+R]	η	Ø	‘lɔvi nɛ:nt	You don’t wash yourself
9	ε	η	tε	η	Ø	η	la	‘lɔvi nɛ:nt	You don’t wash it

10	ACC	DAT	LOC	PRT	
	✓	✓	✓	✓	French, Italian
	✓	✓		✓	Reggio Calabrian dialects
	✓	✓	✓		Friulian
	✓	✓			Spanish, Romanian
	✓				Ardez (Rhaeto-Romance)
					Brigels (Rhaeto-Romance)

The mutually-exclusive properties (1//3-person, singular/plural/unspecified, [±R], [±E]) are used to ‘look-up’ the appropriate (possibly Ø) table entry for each case. As Table 10 shows, Spanish has not developed non-reflexive subject clitics, nor [–SPEC] object-clitics as found elsewhere. As (10) illustrates, availability of clitics varies widely across Romance. We do not pretend that such ‘tables’ exist in any real sense in the human mind, merely that they represent the data in graphically convenient fashion. Nor do we see the properties as traditional ‘features’ available for ‘calculation’. Rather, column and row headings should be seen as classifications, awaiting detailed expression within a wider cross-linguistically adequate semantic/syntactic model. Classifications such as SG~PL are subsets of wider ranging properties (including dual/trial, inclusive~exclusive, mass~count) which are suitable for the divisions active in Romance.²² It follows that there are no uses of feature ‘arithmetic’ in this document. It is our contention that, with the possible exception of 3-3-rules (Chapter 6), the proposed model removes the need for any.

²² We gloss over some distinctions e.g. Amandola (Central Italian, Manzini & Savoia 2005) and Western Ibero-Romance Cantabrian (Fernández-Ordóñez 2009:58-59) ‘re-use’ neuter clitic to distinguish mass~count categories. Along with specificity/definiteness, this area of syntax deserves more detailed study. Here, we simply treat them as 3-person ‘neuters’.

The expansion to four cases leads to simplification. There is no need for clitics to jostle with each other in order to find a place within a limited number of positions, or template; each participant has its own place. An immediate benefit of the $[\pm R]/[\pm E]$ division is that there is no special place for non-active uses of reflexives. Chapter 4 shows that they require no special treatment beyond that already described; non-active anticausative-, middle-, and passive-SE are merely contextually-driven alternatives ($[\pm E]$) of reflexives under NOM/DAT/ACC.

2.2 Against Reductionist Tendencies

Many analysts attempt to reduce duplication of forms by underspecification, driven by notions of ‘simplification’ and/or ‘economy’. In Grimshaw (1997)’s analysis of Italian, 3-person clitics are fully defined, *mi/ti/ci/vi* are only marked for person/number, *si* only for $[\pm R]$,²³ and case is ignored. Following various processes based on these definitions, the full set of properties are added by spell-out rules. Many languages, however, display the exorcised features in their surface forms. Under such an approach, every language has its own active feature set, and underspecification. Similarities between closely-related languages become accidental and cross-linguistic comparison to illuminate shared properties/constructions becomes void. In a case-model, lack of surface-form distinctions between clitics representing clearly different functions does not indicate complex processes of underlying compression, but simple surface-form syncretism. Form and function are distinct.

23 Bruhn de Garavito *et al.* (2002) proposes that *se* is also underspecified for $[\pm R]$.

2.2.1 Functions, *not* Forms

As in many language families, Romance does not show gender on 1/2-pronouns (Kayne 2000).²⁴ From this, it has been argued (e.g. Martín 2012) that such clitics do not carry gender. Since clitics are referents to objects, all their features are readily available; there is no logic for distinguishing any feature as inaccessible. 1/2.SG pronouns refer to individuals whose gender is part of the interlocutors' shared knowledge. 1/2.PL are not 'multiples' of their singulars e.g. *we* does not represent multiple *I*'s, but a group from which *I* is drawn, either excluding (*exclusive-we*) or including (*inclusive-we*) the addressee. Number-marking, therefore, has communicative value, distinguishing individual from group. Gender-marking, however, is superfluous (already known) with no effect on meaning. Moreover, if the gender of speaker/addressee, speaker/group or addressee/group differ, marking is contradictory. It represents added complication without benefit.

Table 12

11	[FR]	Paul a peint les femmes	Paul has painted...the women
12		Paul les a peintes	...them
13	[CA]	En Pere ha pintat les parets	Peter has painted...the walls
14		En Pere les ha pintades	...them
15	[IT]	Mi/ti ha vista _{FEM/OMASC/ONAGR}	He has seen...me/you
16		Ci/vi ha viste _{FEM/IMASC/INOAGGR}	...us/you _{PL}

For French/Catalan, Kayne (2000) notes that while subject agreement expresses number/person (11/13), object agreement on participles displays number/gender (12/14). This extends to optional agreement with 1/2-clitics in several Italian varieties (15-16, Belletti 2001), which must therefore carry gender. There is no reason to assume that dative clitics, also considered genderless, are any different. Some languages show gender on 3.DAT (Italian *gli* vs. *le*, Laísta Spanish *le* vs. *la*), most do not; some show number on 3.DAT (Standard Spanish

²⁴ Spanish/Occitan plural subject pronouns do show gender, but may be bi-morphemic (Martín 2012).

le vs. *les*) others do not (Italian *gli* vs. *gli*). Absence of gender/number-marking on 3-person or 1/2-person clitics is not evidence of underlying absence of the property.

Nor does lack of distinct reflexive~non-reflexive surface-forms, prove lack of underlying [R]. Catalan SE (17) produces three dialect-dependent results in the presence of OBL (18-20, Mascaró & Rigau 2002:11). The expected 2.SG.NOM_[±R] (*te*) may ‘split’ generic reflexive (*se*, same for all persons) from the personal data itself (also *te*, 19), in a process described as “fission” (Halle 2000:132). For some speakers, this leads to the dropping of *te* (person being already indicated on the verb) producing one reflexive form for all persons (20, the “obliteration” process of Arregi & Nevins 2007), however, non-reflexive pronouns e.g. *me*_[±R] never split. This implies that [±R] is present even when not shown distinctly, and *se* is not the only clitic defined for [±R].

Table 13

	N		O	D	A
17		te			
18		te	‘m		
19	se	te			
20	se		‘m		

perds

You get lost (on me)

Similarly, the total underspecification of *se* itself is unjustified. Whilst number does not generally show on *se*, it does in Judeo-Spanish which displays *se*~*sen* e.g. *en biendo*+*sen*, *kozer*+*sen*, =Spanish *Al ver*+*se*, *cocer*+*se* (Penny 2000:180). Similarly, for case. Romanian has unique DAT/ACC forms for each person, whilst Czech uses only *se*_{ACC}~*si*_{DAT} for all persons. Whether case is surface apparent or not, syntactic behaviour is consistent for past participle agreement with ACC but not DAT across all languages (Cinque 1988; Dobrovie-Sorin 1998). See also Schäfer (2008a, 2012a) for syntactic arguments that SE must have case.

Thus, surface-form does not necessarily reflect underlying feature content.

Table 14							
	1	2	3	4	5	6	
-R	me	te	lo	nos	os	los	Judeo-Spanish (Penny 2000)
+R			se			sen	
A	mă	te	se	ne	vă	se	Romanian (Ciucivara 2009)
D	îmi	îți	își	ni (ne)	vi (vă)	își	
A	se						Czech (Naughton 2005)
D	si						

With regard to non-active uses (i.e. as indicators of passive, middle, or anticausative voice), Brazilian Portuguese is particularly illuminating. Whilst the Standard dialect shows the full range of SE usage (30), Vernacular Brazilian Portuguese (Azevedo 1989) shows several variations (31-32). All varieties display true reflexives (21), but drop SE in non-active constructions (22-24). In educated colloquial speech, it is common for 3-person *se* to appear with all reflexive subjects (25-26). Indefinite *se* is infrequent and very rare in speech, except for stereotyped phrases (27). Agent indefiniteness is usually expressed by subject-less 3-person verb forms (28). Although such constructions may be analyzed as deletion of indefinite *se* (29), “there is little reason to suppose such a derivation is part of vernacular speakers’ competence” (Azevedo 1989:866), as research suggests that many speakers are unable to understand constructions with indefinite *se*: “a construção com *se* reflexivo é problemática no dialeto rural não apenas quanto ao uso, mas também quanto à compreensão” (Veado 1982:45). There are, therefore, at least three diastatic clitic lexicons (30-32, somewhat idealised), showing distinct series of clitics for reflexive, non-active, and indefinite uses.

Table 15

21	O Getúlio se matou	G. killed himself
22	...aí um senhor levantou [=levantou-se] para mim sentar [eu me sentar]	...then a gentleman got up for me to sit down
23	O pessoal queixa [=se queixa] muito mas no fim ninguém faz nada	People complain a lot but in the end nobody does anything
24	...depois eu arrependi [=me arrependi] de dizer aquilo	...then I was sorry I said that
25	...eu não é por isso que eu vou se suicidar não [me suicidar]	...that is not a reason for me to kill myself
26	Nós se vemos [=nos vemos] por aí	We'll see each other
27	isso não se diz/faz	One does not say/do that
28	Como fax isso?	How do you do that?
29	Como se fax isso?	How do you do that?

		1	2	3	4	5	6	Indefinite
30	Standard	Reflexive	me	te	se	no(s)	vo(s)	se
		Non-Active	me	te	se	no(s)	vo(s)	se
31	Vernacular	Reflexive	me	te	se	no(s)	vo(s)	se
		Non-Active	Ø					se/Ø
32	Colloquial	Reflexive	se					Ø
		Non-Active	Ø					Ø

Since features are inherited from syntax-tree or referent, there is no benefit to adding further complexity of spell-out/interpretation rules. Such notions of *a priori* simplification or economy, in reality, lead to complexity and inefficiency. In a case-model, clitics are considered fully specified. The fact that syncretism allows some surface-forms to converge (differently in each language) is a separate issue.

2.2.2 Syncretism

Separation of function~form is essential for our understanding of the historical development of these elements. Pescarini (2007)'s study of syncretic forms in Italo-Romance (summarised in 33) shows clearly that whilst some modern forms have converged solely through phonetic erosion (34), most cannot be explained in this fashion. Moreover, contra formalist views, 'aberrant' forms are not arbitrary, but affect particular regions of the clitic lexicon (see column headings of (33) in a systematic, if complex, fashion.

Table 16

33		1.PL			2.PL			3 DAT	PRT	LOC	3 REF
	ACC	DAT	REF	ACC	DAT	REF					
Bologna		s'	s'	s'					(i)n'	i	s'
Sarroch		si	si	si	si	si	si	ddi	ndi	(n)ci	si
Bergamo ²⁵		se	se	se	se	se	se	ghe	ne	ghe	se
Poggio Imperiale		cə	cə	cə	tə	tə	cə	i	nə	cə	cə
Napoli		ce	ce	ce				le	ne	ce	se
Brindisi		nci	nci	nci				nci	nci	nci	si
Bari		ngə	ngə	ngə				ngə	nnə	ngə	sə
Ottanto		nde	nde	nde				nde	nde	nci	si
Campi -LE		nne	nne	nne				nci	nne	nci	si
Catanzarese		nci	ndi	ndi				nci	ndi/a	ci	si
Palermo		ni	ni	ni				ci	nni	ci	si
Lecce		ni	ni	ni				ni	nde	nci	se
Veneto		ne	ne	ne				ghe	ne	ghe	se
Torino		ne	ne	se				je	ne	je	se
Collina d'ora (new)		ma	ma	sa				ga	na	ga	sa
Collina d'ora (old)		ma	ga	sa				ga	na	ga	sa
Vailate		ga	ga	sa				ga	na	ga	sa
Roccasicura		cə	cə	ze				rə	nə	cə	zə
Rocca Imperiale		nə	nə	sə				i	nə	tsə	sə
Arce		ne	ce	ce				glie	ne	ce	se

34	Stage 1	Stage 2	Stage 3	Stage 4
	*nos >	ne >	ne >	ne
	*inde >	nde >	nne >	ne
	Proto-Romance	Lecce-type	Palermo-type	Veneto-type

35	Reanalysis	A	A	B	➡	A	A/B	B	➡	A	B	B
----	------------	---	---	---	---	---	-----	---	---	---	---	---

The early stages of grammaticalization display functional vagueness. The same structure performs two similar functions, which not only acts as a potential motivating factor, but also determines available developmental pathways. Functional re-analysis occurs instantaneously, as a spontaneous activity by individual speakers during communication, as they extend the use of old constructions (and words) to novel contexts. Structural adjustments (structural re-analysis) eventually follow, giving rise to more precise ('iconic') coding of the newer vs. older functions, now as two distinct constructions, allowing them to gradually drift apart

25 SE is used for all persons.

following their own developmental paths. Like biological evolution, structural re-adjustment lags behind functional innovation, and is subject to different constraints and dynamics.

Reanalysis may occur when there are two conceptual spaces with sufficient overlap that one usage may serve for the other in at least some circumstances. With sufficient frequency, learners extract such usages as the target rather than accidental overlap (35). Thus, originally locative *ci*_{PROXIMAL}/*vi*_{MEDIAL} (here, *with us*~there, *with you*) ‘spread’ to replace 1/2.PL no(s)/vo(s) (§5.2.1); reflexes of Latin INCE/IBI > *ci/y/bi/hi* spread from proximal only to all locative uses (i.e. contrastive distal references become generalised place reference); and in many varieties, locatives become impersonal datives, often leading to replacement of the dative in 3-3- or all contexts with the locative form (§6.2.7)).

Faltz (1985) identifies a continuum of reflexive pronominal paradigms from “functionally streamlined” (36) where reflexive forms appear only where ambiguity might arise using non-distinct 3-person markers, to “strategically streamlined” (49) with the same reflexive form for all persons. Various developmental sequences have been proposed e.g. 3/6>4>5>2>1 (Benincà & Poletto 2005), however, Puddu (2010) shows that 3>1>2 and 6>4>5 are also attested. The most robust generalisation is plural>singular, however, data from Milanese and Airolo (de Benito Moreno 2015) indicate syncretism between 3/4/2, without 5. With the (probably accidental) exception of (47), all variants are attested in Romance, often in neighbouring dialects.

Table 17

	1	2	3	4	5	6	Spread of SE to non-III persons ²⁶
36			✓			✓	Surmiran [RR], Orbasque [PI]
37			✓	(✓)	(✓)	✓	Ladin Dolomitan [RR], Castelló [CA]
38			✓		✓	✓	Vallader, Puter [RR], Murcian [SP]
39			✓	(✓)	✓	✓	Ladin Gherdëina [RR]
40			✓	✓		✓	Turinese [PI], Vivaro-Alpine [OC], Mozambican Portuguese
41			✓	✓	(✓)	✓	Bregagliot [LM]
42			✓	✓	✓	✓	Friulian [LM], Picard [FR], Valencian [CA], Río de la Plata [SP]
43		(✓)	✓	(✓)	(✓)	✓	Poschiavino [LM]
44		✓	✓	✓	✓	✓	Medeglia [LM], N. Brazilian Portuguese
45		✓	✓	✓	(✓)	✓	Milanese [LM]
46		(✓)	✓	✓		✓	Airolo [LM]
47	✓		✓	✓	✓	✓	
48	(✓)	✓	✓	✓	(✓)	✓	Mendrisiot, Luganese [LM]
49	✓	✓	✓	✓	✓	✓	Sutsilvan [RR], Bergamasque [LM], Afro-Brazilian Portuguese

[CA]=Catalan, [FR]=French, [LM]=Lombard, [OC]=Occitan, [PI]=Piedmontese, [RR]=Rhaeto-Romance, [SP]=Spanish

Se not only ‘spreads’ within ‘reflexive’ paradigms, but across paradigms e.g. Ladin Dolomitan *se* moves into ACC_[−R] (but not DAT_[−R]). Conversely, in Eastern Peninsula Spanish including Valencia and Murcia, M.PL.ACC *los* can be used as 4/5 object clitics [±R] (17, Enrique-Arias 2011). Rumantsch varieties Surmiran and Surselvan show distinctions between reflexive and non-reflexive clitics in 1/2.SG, in stark contrast to most other languages.

Table 18

	1	2	3	4	5	6	
−R			lo	nos	os	los	Standard Spanish
+R	me	te	se			se	
−R			lo	nos	os	los	Judeo-Spanish (Penny 2000)
+R	me	te	se			sen	
−R			lo/a	nos/los	os/los	los/as	Murcian Spanish (Ordóñez 2002)
+R	me	te	se			se	
D			(t)i	nes	ves	(ti)	Ladin Dolomitan (Meneghin 2008)
A	me	te	l/a	nes/se	ves/se	i/les	
R			se	nes (se)	ves (se)	se	
−R	am	at	igl/la	ans	az	igls/las	Surmiran (Anderson 2016)
+R	ma	ta	sa			sa	

26 Parentheses indicate that SE, the expected pronoun, or a combination may be used. de Benito Moreno (2015) for more dialects, examples and references therein.

Such variation of development can only occur if form is a separate property from featural make-up. Such cases underline the need to study clitics in terms of the functions which they perform (an indication of underlying features) separately from their surface form.

2.2.3 Null Entries

The number of empty entries in Table 10 (p.42) might cause surprise. This is in part because Spanish lacks adverbial forms, but this doesn't mean that it lacks such clitics, merely that they are \emptyset . Empty slots are meaningful: there is as much contrast between *me~\emptyset* as *me~te*.

Table 19		
50	–¿Qué has hecho en los últimos años?	What have you done in the last years?
51	–He enseñado \emptyset_{ACC} , como siempre	I have taught ($\emptyset=matemáticas$), as always
52	–Nada, he donado [mi tiempo] \emptyset_{DAT}	Nothing, I have donated my time ($\emptyset=a la gente$)
53	He donado [mi tiempo] [a programas de beneficio social]	I have given my time to social programs
54	\emptyset_i hablaron durante muchas horas	They spoke (words) for many hours
55	¿ \emptyset_i hablas Inglés?	Do you speak English _i ?
56	{ $\sqrt{lo}/*\emptyset$ } veo a Juan	I see Juan
57	No { $*\sqrt{lo}/\emptyset$ } veo a nadie	I see no-one
58	A: ¿Juan se compró vestidos _i ? B: Sí, se compró \emptyset_i	Did Juan buy clothes for himself? Yes, he bought himself some
59	Los sapos \emptyset_i repugnan e_i .	Toads are repugnant to everyone_i
60	Los sapos le_i repugnan a [todo el mundo _i]	Toads are repugnant to everyone _i

Usually, valence object arguments are obligatorily filled, however, in order to produce generic statements, either may be omitted, implying abstract theme (51)²⁷ recipient (52). Specific arguments, however, must be overt (53). Similarly, ‘inherent’ accusatives, where verbs have lexicalized their object within their meaning (Talmy 1985), are simply [–SPEC] arguments lexically licensed by the verb (54), which may be ‘over-written’ by [+SPEC] arguments (55). In accusative-doubling the clitic must agree with its [+SPEC] (56) or [–SPEC] (57) referent, as it must when used anaphorically (58). In (59), lack of an experiencer (as seen in 60), makes the statement more generic, by highlighting the repugnance as a property of the toads rather

27 cf. French *Le problème n'est toujours pas résolu, mais j'écirai __ au ministère* (Melis 2004:172).

than a reaction of people. In case-models, such sentences are *not* seen as argument omission, but rather filling argument slots with clitics representing [–SPEC,DAT], [–SPEC,ACC], [–SPEC,OBL] which happen to be Ø.

Similarly where languages such as Spanish lack adverbial clitics. §5.5.6 shows that *ne*_{ABL} in Italian *andarsene* is present in its Spanish equivalent, merely represented as Ø. In some cases, the ‘missing’ forms are not Ø: see *le*_{DAT} for *ne*_{GEN} (§5.2.2); i.e. surface forms may be ‘lost’ by another form ‘spreading’ to its position in the clitic lexicon. Equally, loss of Spanish locative *y* (XV^c) is associated with wider changes such that ditransitive indirect-object *a*-NPs are now read, by default, as essentially locative, with ‘doubling’ clitics forcing dative-recipient readings (§3.2.5), i.e. clitic~Ø has become meaningful in its own right. The need for overt forms is determined by language-wide contrast. Northern Italian dialects have NOM-clitics, whilst most Romance languages, these always surface as Ø.

The existence of null clitics also leads to natural explanations of many ‘random’ exclusions as simple agreement e.g. Spanish *SE*_{IMP} cannot take reflexives because its [–SPEC] object-clitics are defined as Ø; unlike Italian which has such forms resulting in *ci+si* (§4.6.9). Similarly, 3-3-processes follow from simple agreement and look-up; it is merely that in these cases the entries arrived at are generally not Ø, but filled by a surface-form which may also be used in other circumstances (§6.2.7), engendering ill-defined processes such as the spurious-*se* rule.

Far from introducing unwarranted complication, positing empty slots actually makes the comparison of languages more coherent and simplifies each language’s grammar. Speakers know which clitics surface overtly and which are realized as Ø. If Ø contrasts with overt

clitics in the same position/context, pronominalization as Ø will be communicationally meaningful, otherwise alternative constructions are used.

2.2.4 Unrealistic Expectations

Whilst it would be convenient for analysts if clitics took different forms in each position/function, any *expectation* that this should or could be so, ignores the nature of the object under consideration. Surface-form convergence is the natural result of Latin's initial limitations and vicissitudes of phonological development. Indeed, it is effectively required during conversion from WPs to clitics, since the latter are by their nature prosodically reduced and hence unable to carry much phonological information.

The inherited initial consonants *m/t/l/n/v/s* carry most of the important number/person-identifying information. Only by introducing further (and historically unsourced) consonant bases could matters be made more explicit. Whilst some new forms did result (e.g. Italian *ci*, Old Spanish *ge*) from natural phonological changes, most languages have tended to reduce their phonological range even where this collapses distinctions e.g. loss of Spanish palatal consonants saw Old Spanish *ge* [ʒe] > *se* [se], even though the result is identical with existing *se*, i.e. introducing real surface ambiguity.

The potential for distinguishing vowels is also limited. Rapid succession of unstressed monosyllables does not lend itself to strong distinctions being maintained. Languages tend to select default vowels (Spanish *e*, Italian *i*) which merely serve to separate out the information-carrying consonants, whilst allowing phonological processes to apply which further reduce distinctions e.g. Italian *mi* → *me* ___ *ne/lo*. i.e. losing dative~accusative distinctions. Only for

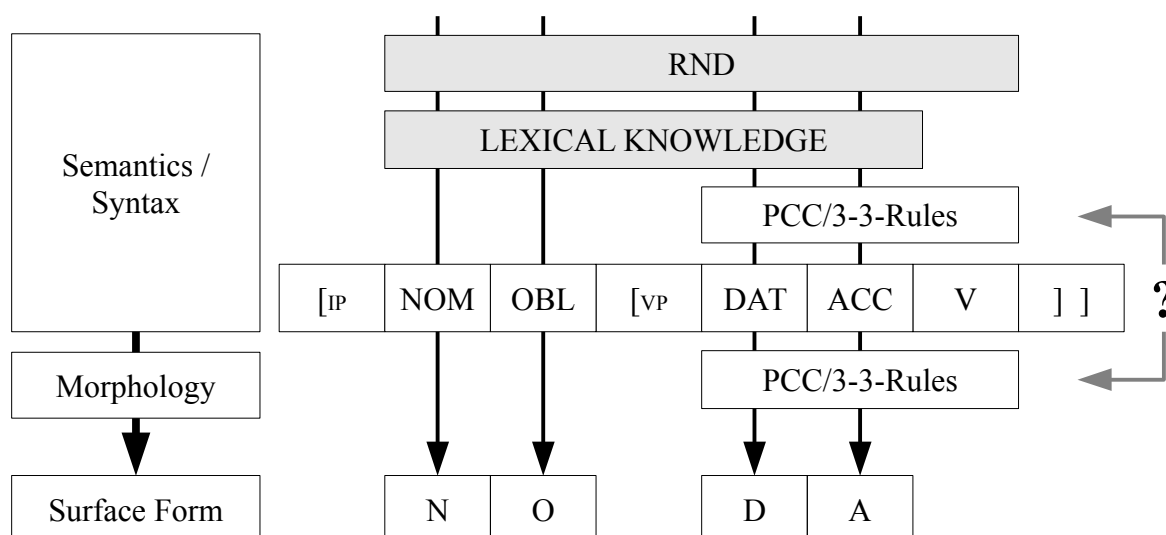
3.ACC where disambiguation is crucial (Spanish *lo/la/los/las colpó*, ‘I hit him/her/the men/the ladies’) is any significant distinction made, and noticeably this is at the end of the phonological sequence where distinctions are easier to hear and maintain. The fact that no Romance or Slavic language has sought to force such distinctions in itself indicates that it is not necessary for effective communication. Indeed, if such distinctiveness had been necessary, these pronouns would not have developed into clitics.

Languages maintain forms in ways which reflect real needs for distinction: greater variation where needed, less where it is not i.e. true ‘economy’. For 1/2-persons, there is only one possible anaphoric referent for which the listener already knows its gender, and whether it is reflexive by virtue of verb ending. There is no need to mark this by different surface forms, but that does not mean that the syntax/morphology is unaware of this data. Only in the 3-person is there room for doubt (since there may be more than one 3-person referent) and here, there is more surface distinction. Certainly more forms would be useful out of context, however, clitics are the glue that holds discourse together; they can’t be removed from context, and context offers all that is required to make the necessary inferences.

In short, analysts should not expect explicitness of surface-form. Hence, lack of explicitness is not an argument for lack of underlying specification. Rather, every clitic is an expression of the combinations of features from the syntax and its referents, which is ‘looked-up’ in the table to find its historically-arrived-at surface-form. Syncretism with another feature-combination is irrelevant. It merely means that analysts have to look more carefully at how such surface-similar forms can be parsed into different underlying structures.

2.3 Exclusions

In our case-model, failure of particular sequences to surface is not related to template-sequencing nor person-hierarchies,²⁸ but based on ‘exclusions’ of various types, which contra MC models, are inherent in the structure and not post-lexical after-thoughts.



In our proposal, clitic combinations are restricted by the PCC (limited to the lower clitic-field), RND (operative across the whole clause) and knowledge of the clitic lexicon (lexical knowledge, operating across the entire language). At this point, we remain agnostic to where the PCC operates: semantics, syntax or morphology. That issue is developed Chapter 7.

2.3.1 RND

Person-models depend upon numerous surface exclusions e.g. **me+te*, **le+lo*, implemented as surface-form constraints, morphological feature operations, and/or person-hierarchies. Whilst such rules *describe* situations, they lack explanatory power.

²⁸ Addressed in Chapter 7.

We start by isolating cases relating to the same grammatical person e.g. **me+me*, **me+nos*, explained by Strozer (1976) in terms of exact vs. intersecting identity. Crucially, such restrictions hold not only between clitics, but also between clitic and verb (**Nosotros me salpicamos* ‘we splashed me’), and hence are beyond the reach of putative MCs.

	1	2	3	4	5	6		1	2	3	1	2	3
1								1	+R				
2								2		+R			
3								3					
4					??	?		4			+R	??	?
5				??		?		5			??	+R	?
6				?	?			6				?	
DAT to ACC Clitic								Clitic to Verb Subject					
	Exact Identity								Distinct Referents				
	Intersecting Identity							?	Potential Overlap				

We propose a language-wide restriction of **Referent Non-Duplication** (RND). Syntactically, this follows from the observation that once a referent’s ϕ -features have been absorbed in one position, they are no longer available to other positions.²⁹ §7.5.5 presents RND as a semantic restriction reflected in, and expressed through, structure.

RND excludes cases of exact and intersecting references between clitic~clitic and clitic~verb, which incorporates a limitation to one reflexive per clause. As indicated, RND disallows cases of two 3-person clitics with overlapping referents, but allows pairs with disjoint referents. This last property has important consequences for the nature of 3-3-mutations (Chapter 6).

²⁹ Cf. Laenzlinger (1993)’s Principle 4, “Two clitics with the same referential value for individuation cannot co-occur within the same derivational domain”, or Chomsky (1981:36)’s theta criterion: “Each argument bears one and only one θ -role, and each θ -role is assigned to one and only one argument”.

Studies such as Evans *et al.* (1978) and Lepschy & Lepschy (1984) show that these combinations are always *ungrammatical*, whereas *acceptability judgements* for remaining combinations vary between surveys, languages, informants, and even for the same cluster in different contexts. This leaves the remaining exclusions (including the PCC) as a further filter over and above strictly *grammatical* (i.e. syntactically deviant) restrictions controlled by RND, allowing us to capture those properties which are shared by all Romance languages whilst highlighting those areas which may be language-specific.

The diagram also highlights potential difficulty with multiple plural referents which has been used as support for number-based morphological processes. In Chapter 7, we show that the acceptability of combinations such as *nos*_{1,PL}+*os*_{2,PL} varies with speaker perception of the situation: it is considered to be acceptable if referents are clearly isolatable, but unacceptable if they imply overlapping e.g. *we+you* implies a 'greater' *we*. We propose that such constraints should be seen as part of the proposed semantic restraints (RND), rather than discrete morphological processes.

2.3.2 PCC

Introducing NOM/OBL has the effect of moving many clitics out of the PCC's control leading, along with the approach to exclusions, to a simpler definition of the PCC itself. Defining the clitic-field in terms of two sub-fields, also allows us to delimit its space of operation. §7.5.3 shows that there are no operations for putative MCs to perform in the upper field, thus limiting any MC to DAT/ACC, e.g. **me*_{DAT}+*te*_{ACC} and not **me+te*, which is legitimate in other circumstances.

Once differences based on availability of clitics (lexical knowledge) have been abstracted, general syntactic exclusions (RND), and simple mutual exclusion within the same node, what remains to the PCC is a simple set of exclusions e.g. Spanish **me_{DAT}+te_{DAT}*, **te_{DAT}+me_{ACC}*, **le_{DAT}+me_{ACC}*, **le_{DAT}+te_{ACC}*. Contrary to many previous proposals, there is no justification to consider these in terms of person-ordering; they are merely exclusions, which may broadly be described as ‘[+human,ACC] entities may not be possessed by [±human,DAT]’. §7.5.5 shows that these emerge naturally from syntactic structure for HAVE-languages e.g. Spanish, but are only partially applicable in BE_AT-languages, thereby explaining the different behaviour between Romanian and the rest of Romance with respect to the PCC. Finally, §7.4 shows that putative PCC breaches are in fact merely the use of existing functionality which ‘look like’ the usually excluded combinations.

In our opinion, the significance of the source of RND’s restrictions has generally been overlooked. Constraints such as **me+me/*me+nos/*nos+os* are outside of putative MCs and clitic-specific syntax. From our perspective, the fact that so many ‘exclusions’ *cannot* be part of a morpho-syntactic exclusion mechanism should engender a certain scepticism concerning all exclusions. Chapter 7 looks at a way of removing the very concept from the model.

2.4 Conclusions to the Model

Despite being a very simple model, we contend that it is capable of meeting all communicative needs. In fact, in our opinion, it is due to being so simple that this is possible.

The following chapters ‘fill in’ the boxes in our clitic-lexicon tables: DAT~OBL (Chapter 3), reflexive and non-active SE (Chapter 4), and non-personal clitics (Chapter 5). In these

chapters, we hope to show that an *iconic* structure allows speakers to *compose* and listeners to *interpret* messages *in context* without confusion, regardless of surface similarity, and without the need for lexicalization of “unanalyzable chunks” or complex mechanisms to control surface order. The last chapters turn to the effects of the model which, we argue, are to remove most (possibly all) need for inter-clitic manipulation (Chapter 6) and pattern restrictions (Chapter 7).

3 TWO DATIVES

This chapter explores personal indirect clitics which we divide between DAT and OBL (§3.1.1) reflecting the upper vs. lower clitic-field division. We introduce the central concept of $[\pm E]$ (§3.1.2) which permeates all following chapters representing the key distinction between coincidence (disjoint reference) vs. possession (subset reference). In addition to outlining the first tranche of the proposed structural model, the chapter discusses the need for inference as an inherent part of the nature of language, which we see as supported by that structure. Once presented in this ‘accessible’ scenario, we will be ready to apply ‘case’ and $[\pm E]$ to the more complex areas of reflexive (Chapter 4) and non-personal (Chapter 5) clitics.

3.1.1 OBL~DAT

In addition to prototypical ‘transfer constructions’ with person/place goals, ‘datives’ often perform functions unrelated to verbal valency, ranging from ‘inactive agent’ (1) to discourse emphatic of politeness (6). This variety has proven difficult to express in a coherent motivated model, resulting in multiple classificatory systems, and conflicting terminology. Our model defines two classes of datives, where the level and type of *affectedness* represented by each class reflects the clitic’s structural position (DAT vs. OBL³⁰) each of which possesses the property $[\pm E]$ described below.

Table 20

1	Le encantó la película	He loved the movie
2	Le arruinó la fiesta a Valeria	He ruined Valeria’s party
3	Me le arruinó la fiesta a Valeria	He ruined Valeria’s party on me
4	El problema se me fue de las manos	The problem escaped from my hands
5	Se leyó el periódico de una sentada	He read the newspaper in single sitting
6	Pásele!	Come on in!

30 All languages surveyed in Polinsky (2005) make use of affected ‘experiencer’ functions, over half use locative/instrumental functions, whilst comitative/substitutive functions are common with intransitives.

The distinction between indirect-object (DAT) and ‘other’ datives (with various names) has long been recognised e.g. DAT clitics are PCC-controlled and their absence changes sentence meaning/grammaticality, but ‘other’ datives introduce participants free from the PCC with no effect on grammaticality (Perlmutter 1971; Morin 1979; Albizu 1997; Ormazabal & Romero 2007; Bianchi 2006). Despite such clear differences, OBL is never treated on a par with other ‘cases’. Whilst DAT is seen as something concrete, OBL (when considered at all) is vague and additional. This chapter focuses upon the need for, and benefits of, recognising two types of semantically *and positionally* distinct ‘datives’.

3.1.2 [±E]

Since Benveniste (1966a) treating possession as an inclusive locative relationship, where HAVE=BE+Preposition, has been widely exploited. Urban dialects of Palestinian Arabic (Boneh & Sichel 2010) possess BE, but keep the ingredients of HAVE separate, overtly distinguishing part-whole and coincidence by choice of preposition.

Table 21

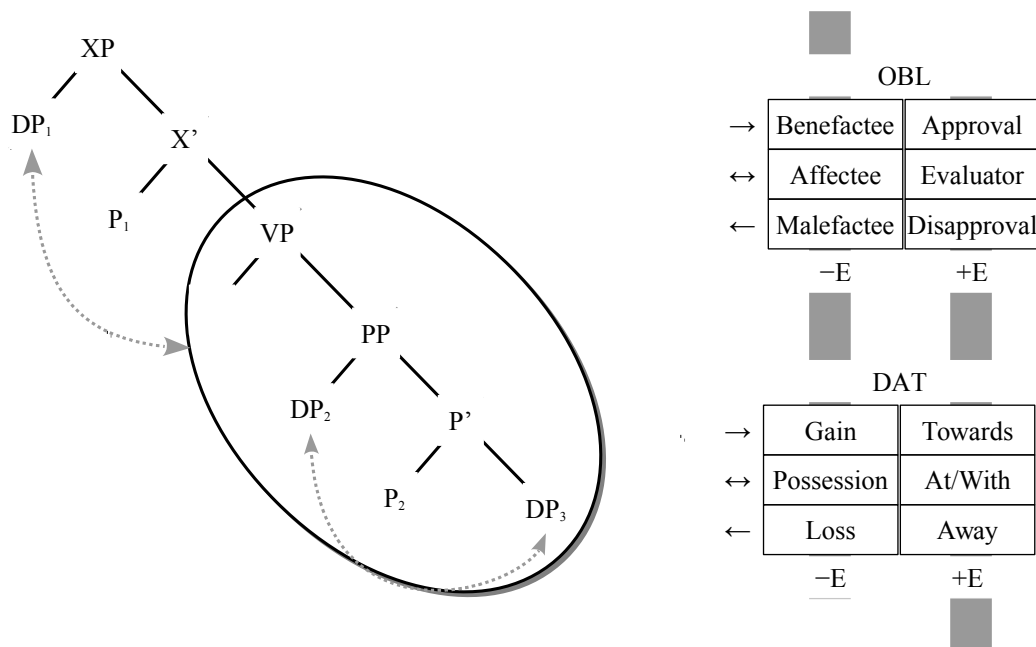
	(a) [−ANIM] Possessor		(b) [+ANIM] Possessor
7	kaan la-əš-šajara ʔruʔ ktar WAS.3SG.M to-the-tree branches many The tree had many branches	[−E]	kaan la-mona ʔanf t awil/tlat ulaad WAS.3SG.M to-Mona nose big /three kids M. had a big nose/three kids (as a mother)
8	kaan ʔind əš-šajara ʔruʔ ktar WAS.3SG.M at the-tree branches many Near the tree were many branches	[+E]	kaan ʔind mona ktaab WAS.3SG.M at-Mona book M. had a book
9			kaan la-mona lat ulaad #kull yom WAS.3SG.M to-Mona three kids #every day Mona had three kids #every day
10			kaan ʔind mona tlat ulaad WAS.3SG.M at-Mona three kids M. had three kids (as babysitter, possibly mother)

With inanimate NPs, *la-* marks part-whole relations (7a), whilst coincidence is marked by various locative prepositions (8a). With human possessors, body parts and kinship³¹ are

31 Inclusion of social relations and kinship in part-whole/inalienable relations is language-dependent (Baron *et al.* 2001; Heine 1997).

indicated by *la-* (7b), and looser associations by locative preposition *find* ‘at’ (8b). In some contexts, *find* may imply kinship. In (10b), *Mona* is a babysitter, but kinship may be inferred. In (9b), *la-* forces a part-whole relationship reading i.e. motherhood.

We represent this key division as external ([+E], ≈coincidence) vs. internal ([−E], ≈part-whole), leaving definitions somewhat abstract, merely an opposition. This is necessary since, as we will show, clitics solely indicate the presence of relationships, where their ‘meaning’ depends on the items being related and the context within which the relationship is defined. Most details are inferred from context and world knowledge e.g. personal clitics do not indicate ‘direction’ (i.e. ‘to’ or ‘from’ a possessor) which must be inferred from verbal semantics/situation. This is not ambiguity, but vagueness: when significant, arguments appear as PREP+Complement.



We define a similar relationship between OBL and VP. Due to the difference in the nature of subordinate partners, interpretations (although clearly related) also differ. Clitics indicate

presence/absence of secondary participants relating to ACC or VP. The rest is inference.

Although the current model does not make use of any specific theoretical apparatus other than the existence of functional heads in IP, it does map quite closely to the concepts of applicatives. Such heads are divided between ‘low’ and ‘high’ (Pylkkänen 2002); ‘entity-related’ and ‘event-related’ (Cuervo 2003). According to Harley (1995, 1998, 2002), Cuervo (2003) and McIntyre (2006) *i.a.*, applicative heads have very reduced semantics, merely establishing an abstract HAVE-relation between specifier and complement. The exact interpretation derives from the type of structure to which it is applied, and the availability of such constructions in each language. Our model uses $[\pm E]$ to differentiate the HAVE-relationship between ‘possession’ and ‘coincidence’.

3.1.3 Patterns Available

(11-22) introduce uses of, and restrictions upon, ACC/DAT clitics with some examples of OBL to illustrate relative position and lack of person restrictions (§7.5 for all permutations). Since OBL has no direct English equivalent, the phrase ‘on X’ is used in the translations. This can sound awkward, although similar usages exist e.g. *They did the dirty on him*.

Monotransitives introduce *effectees* which may be substituted by clitics agreeing in number/gender (11). Ditransitives introduce a further *affectee*. Whilst dative case (with separate forms) has survived in Romanian (24), it is represented by PPs elsewhere in Romance. Thus [a Pablo]_{DAT} (12) acts as a unit indicating dative case, which may indicate source (14) or destination (13) of ACC. DAT clitics are $[-E]$, i.e. they cannot be used as $[+E]$ locatives (15~16).³² Contra many analyses, Cuervo (2003) notes that DAT is *not* restricted to

32 §7.4.4 shows that Romanian is, once again, an exception to this observation.

humans, some inanimates (17), but not all (18) can ‘possess’. A static relationship is indicated by application of datives to monotransitives (19). In these cases, ACC cannot be pronominalized, nor possession duplicated e.g. [de Pablo]_{GEN} (20-21). Finally, benefactives introduce ‘intended’ goals either through PP or clitic (22). Contra most Romance languages where clitic and referent are mutually exclusive, Spanish and Romanian allow DAT to be doubled (15), except for cases of static possession.

Table 22

	Topic/S _H	O	D	A				
11				<la _j >	come _i <la paella _j >	He _i eats _i {it _i /the paella _i }		
12			Ø _k	lo _j	da _i [a Pablo] _k	He _i gives _i it _j [to Pablo] _k .		
13		(me) _i	te _k	lo _j	da _i	He _i gives _i it _j to you _k (on me _i).		
14		(te) _i	me _k	lo _j	roba _i	He _i steals _i it _j from me _k (on you _i).		
15			(le _k)		mandó _i un libro _j	a Gabi _k	He _i sent _i a book _j	to Gabi _k .
16			*le _k			a Barcelona _k		to Barcelona _k .
17			le _k		puso _i azúcar _j	al café _k	I _i put _i sugar _j	in the coffee _k .
18			*le _k			a la mesa _k		on the table _k .
19			le _k		lavó _i la bicicleta _j	[a Pablo] _k	He _i washed _i Pablo's _k bicycle _j (on me _i)	
20		(me) _i				*[de Pablo] _k		
21			le _k		lavó _i (*su) bicicleta _j		He _i washed his _k bike	
22		<le _k >			hace _i la torta _j	<para él _k >	He _i makes _i the cake _j for him _k (another)	
23			se _i	Ø _j	lava _i las manos _j	He _i washed _i his _i hands _j		
24			și _i		a _{AUX} spălat _i mîini _j +le _{DEF,ART} ([RO])			
25			s' _i		a _{AUX} spălat _i pe _{PREP} mîini _{ACC} ([RO])	They _i criticised		
26			se _i		critican _i [a los mismos _i]			
27			se _i		critican _i [los unos _i a los otros _j]	each other _i		
28	A Pablo _k ,	le _k			gustan _i los libros _i	Books _i are enjoyed _i by P _k .		
29					M. ya _{PAST} -tambu _{WALK} -le _{APPL} -dde _{PAST} K.	Mukasa walked for Katonga.		
30	Hugo _i	*le _k			corrió _i a Vicki _k	*Hugo _i ran _i for Vicki _k .		
31		le _k		Ø _j	corrió _i una carrera _j a Vicki _k	Hugo _i ran _i a race _j on Vicki _k .		
32	Juanita _i ya	le _k			camina _i	Juanita _i already walks _i on him/her _k .		

Subject coreferent objects take reflexive forms of the appropriate case, with the same limitations on possessor datives (19). The distinction is clear in Romanian (*și*_{DAT}~*se*_{ACC}); in (24), the subject is possessor of, in (25) he is, the object. The same relationship holds for transitive-reflexives (26) vs. reciprocals (27).

OBL may appear with intransitives (28) acting as event experiencer. Spanish OBL does not employ the full range of possibilities found across languages; compare Lugandan (29, Pylkkänen 2002:25) vs. (30). Similar sentences are acceptable when verbs are transitivized (31), or where the experience can be related to the event as whole (32).

Table 23

33	Luca mi pedala male	Intransitive/unergative	Italian
34	Luca mi è caduto	Intransitive/unaccusative	
35	Luca mi si è ammalato	Intransitive/middle ('pronominal')	
36	Luca mi mangia troppo	Transitive	
37	Luca mi ha dato la lettera a Maria	Ditransitive	
38	Lucia mi si mangia una mela	Indirect reflexive (benefactive)	
39	Lucia mi si mangia le unghie	Indirect reflexive (possession)	
40	Lucia non mi si lava	Direct reflexive	
41	Tua madre mi gli fece scrivere la lettera	Your mother made me write him a letter	
42	Mi gli scrivi queste lettere?	Would you write him this letter for me?	

Similar patterns are found across Romance, with some differences in usage e.g. availability of SE_{ANT} is more restricted in French/Italian than Spanish (§3.3.5), hence $SE_{ANT}+OBL$ is more frequent in Spanish, where just OBL is used in French/Italian. There appear to be no other restrictions, with OBL applying to all verb types (33-40). Clusters of two 'datives' are common when they were originally governed by different predicates (clitic climbing, 41), or if one is an event benefactor (42). Data from Lepschy & Lepschy (1984:213).

3.1.4 Chapter Outline

This chapter focuses on Spanish as displaying the greatest freedom in its use of both DAT and OBL. §3.2 investigates DAT finding that empirical data does not support the hard and fast rules usually presented for it. §3.3 discusses OBL showing that its use is no less clear than DAT and is best expressed by a separate position. §3.4 considers areas where interpretation of the two fields may appear to overlap. In fact, a clear understanding of the OBL~DAT divide provides answers to many previously difficult questions. We argue that, not only in order to

include OBL, but also to explain real-life use of DAT, a more abstract view of clitics is required where both are vague, and never directly translatable, but rather signal significant relationships *within* (DAT), or *relating to* (OBL), the event. ‘Meaning’ can only be inferred (§3.5) from context and, if both are present, contrasted by position. Only by understanding the balance between both types of datives can either be understood. Only by separating them out positionally can real-world data be accommodated.

3.2 Lower Clitic-Field

For most Romance languages, DAT_[-R,+E] clitics are Ø, so that only ‘possession’, not ‘coincidence’ *within* the event, can be expressed through clitics. Romanian does possess DAT_[-R,+E] giving it relative freedom from the PCC, as discussed in (§7.4.4). In addition, most have OBL_[±E] clitics capable of indicating ‘possessive’ and ‘coincidence’ *with* (the effects of) the event, although ‘coincidence’ (OBL_[+E]) paradigms are often restricted (§3.3).

In the lower clitic-field, the key relationship is between DAT and ACC, usually described in terms of ‘possession’. This is a useful term used throughout the work, but cannot be understood as ‘possessor raising’ with specific rules for its (non-)appearance, as usually presented in grammars. Use of DAT clitics *requires* interpretation, which may include part-whole relationships, possession, ownership, each of which may be seen as a specific examples of a far looser link, better described as *affectedness*.

3.2.1 A Note on Translations

English glosses mask significant differences with Romance. Spanish can express possession/ownership through possessive adjectives, but tends not to do so where ownership is ‘obvious’ (43). Spanish defaults to readings of subject possession; *la* implies *su* (43), whilst

su requires particular justification (44). In English, which expects possessive adjectives, *the* defaults to readings of external possession, leaving listeners searching for someone-else in the context to act as possessor.

Table 24

43

Levantó_i (la mano)_j

(?)He_i raised_i the hand_j → He_i raised_i his hand_j

44

(?)Levantó_i (su_k mano)_j

He_i raised_i his hand_j.

Default

Specific

Contrastive

Spanish

Possessor of j

English

Possessor of j

la

i

his

i

la

i or k

the

k

la

k

the

k

su

i

his

i

Readings of external possession are acceptable in both languages in specific contexts e.g. a mortician raises *the* hand (of a cadaver). This meets English expectations, whilst requiring no change (and providing no greater clarity) in Spanish. Contrastive situations (e.g. a mortician with his own and someone-else’s severed hand before him) may be clarified by introducing the unusual *the* (43) or *su* (44). *Su* is avoided, therefore, not due to its *ungrammaticality* but rather to its unnecessarily emphatic quality, implying something beyond the norm, and leaving Spanish listeners searching context for someone-else as possessor such that this specificity is necessary.³³ Thus, English and Spanish have opposite default readings for possession. Whilst the translation ‘his’ is often appropriate/necessary where it is absent in the Spanish, there is no one-to-one correspondence.

3.2.2 ‘Dative’ ≠ Possession

Spanish may also express possession through DAT clitics (45-46), which are putatively obligatory when subject coreferent (47).

33 Similarly, subject pronouns are obligatory in English carrying no semantic weight, but unnecessary in Spanish, where their use is restricted to emphatic situations; use in ordinary situations communicates something extra which is inappropriate to the situation.

Table 25

	N	O	D	A		Possession		
45			le _k	Ø _j	cortaron _i (la mano) _j .	They _i cut _i off his hand _j .	External	
46							They _i cut _i off his hand _j .	Internal
47			me _i	Ø _j	corté _i (la mano) _j .	I _i cut _i (off) my hand _j .	Subject	
48						*/?corté _i (mi mano) _j .	I _i cut _i my hand _j .	Subject
49						?corté _i (su mano) _j .	I _i cut _i his hand _j .	External
50			le _k	Ø _j	mandó _i el hijo _j .	He _i sent _i his _{i/l/k} son to him _j .	Any	

If we gloss *cortar* as *cut off*, (45) is ditransitive (*≈remove*) with *le* realizing the source from which possession is lost. A gloss of *cut* (46), however, where the hand remains with its owner (monotransitive like *levantar*, 43-44) is also possible. In all cases (45-47), dative clitic usage remains the norm, despite the fact that (as shown above) there is no requirement to indicate such possession in cases of co-reference, and only for purposes of clarification in external possession. Furthermore, as with *levantar*, cases of questionable acceptability (48-49) may be felicitous in context, thereby refuting the obligatory nature of the rule. Finally, ditransitives pose the opposite problem where three readings of possession are possible according to context (50). Analyses of DAT directly as ‘possession of ACC’ are, thus, incoherent.

3.2.3 ‘Dative’=Affectedness

DAT’s primary function is *not* to express possession, but rather *involvement within the event* from which possession/ownership may be inferred. This is evidenced in cases where DAT cannot be used when possession is true, can be added where it is incorrect, or removed where it might be expected. Examples from Tuggy (1985).

Table 26

	O	D	A					O	P	A
51		*le _k					*his _k son _i	+	+	–
52		me _k	Ø _j	vieron _i	al hijo _j	They _i saw _i	my _k legs _j (muslim lady)	+	+	+
53		me _k			las piernas _j		my _k books _j (dishonest accountant)	±	+	+
					los libros _j					
54		Ø _k			su _k coche _j	They _i got _i his _k car _j dirty		+	–	–
55		le _k	Ø _j	ensuciaron _i	el coche _j			+	±	+
56	le _k				el coche _j	They _i got _i the car _j dirty on him _k		±	±	+
57	le _i	Ø _k			tu _k coche _j	They _i got _i your _k car _j dirty on him _i		–	±	+

In (51), a father (*le*) is not affected by the event of his son being seen; in this case, a clitic is considered *ungrammatical*. However, *being seen* can *affect*, whether possession is inalienable (52) or not (53). In each case, possession is *inferred* as cause of the affectedness. In (54), ownership is declared by *su*, but the owner is construed as unaffected, or irrelevantly so. In (55), ownership (and possibly possession) is inferred from *his* being affected (indicated by *le*). However, (56) provides an alternative reading, where ownership/possession may or may not be true, but affectedness remains. The correct reading is derived from context; *not* surface form. In (57), ownership is specifically denied; *he* is affected because *he* is responsible for looking after *your* car, regardless of whether the car was in *his/your* possession. As indicated, ownership [±O] and/or possession [±P] vary; only *affectedness* [±A] is constant.

Table 27

		D	A					O	P	A
58		le _k	Ø _i	abrieron _i	el estómago _j	They _i opened _i his _k stomach _j	[±conscious]	+	+	+
59		Ø _k			su _k estómago _j		[–conscious]	+	+	–
60		me _i	Ø _j	el diente _j			To dentist, before extraction.	+	+	+
61							Displaying it, after extraction.	+	+	–
62	Mire _i		Ø _j	mi _i diente _j		Look _i at my _i tooth _j .	To analyst to whom the tooth has been sent.	+	–	–
63			Ø _j	el diente _j			Discussion of an independent tooth.	–	±	–

Conversely, affectedness may be denied by removing the clitic in order to highlight lack of awareness (58-59) or physical alienation (60-63). In (60), the participant is affected by pain (cf. *me duele el diente*) caused by possession, but not after its removal (61). In both cases, *he* is possessor and owner. In (62), *he* remains the owner, but no longer possesses it, whilst (63) indicates that no-one is affected by possession/ownership. The ‘obligatory’ nature of coreferent clitics is because, in most situations, subjects are affected by ownership/possession, but there is no ‘rule’ enforcing this and, therefore, no ‘exceptions’ to it. Absence cannot be considered a ‘rule’ exception.

Table 28

		D	A		
64			la _i	irritaba el roce de la cinta	The rubbing of the tape irritated her _i
65	A ella _i	le _i		irritan mis atenciones	My affections irritated her _i
66	Los perros		lo _i	molestan siempre que llega ebrio	The dogs harass him _i whenever he arrives drunk
67		le _i		molestan (*siempre...)	The dogs bother him _i (...in general)
68			se _i	bañó	He bathed (himself _i)
69		te _i	Ø _x	enfadas _i	You _i are getting annoyed _i

A large set of verbs may appear with accusative or dative, translated by identical (64-65) or different (66-67) lexemes depending on the receiving language (Vázquez Rozas 2006 for lengthy lists). Physical effectedness tends to accusative (64), whereas psychic affectedness tends to dative (65, Hurst 1951:76). Ackerman & Moore (1999:9), following Treviño (1992), contrast ‘direct affectedness’ and ‘non-direct affectedness’ (66-67). In our terms, (66)’s participant is *effected* as the object, whilst (67)’s object is the inherent Ø_{ACC} (e.g. feelings) which undergoes a change-of-state *affecting* their possessor (DAT). In (68, ACC), there is no sense of affectedness, whilst ‘inherent’ reflexives (69, DAT) show pure affectedness by virtue of possession of an inherent ACC (e.g. sensibilities). See Chapter 4 for use with reflexives.

3.2.4 (In)alienable Possession

Signalling possession through dative clitics is unacceptable where possession is expected/inalienable (70). Presence of a dative clitic is appropriate when interacting with body parts as external items (71), and required when the subject uses their hands as external instruments (72). The pattern extends to alienable objects, where SE indicates that such objects are considered part of the subject's dominion (73-74). Without SE, actions are performed for another participant, represented by [-SPEC,DAT] clitic \emptyset_x . Kliffer (1983) notes that (75) may have (in)alienable readings: by default, the skirt is considered to be Mariana's, however, when trying it on in a shop, possessive interpretations do not obtain.

Table 29

70	Tenía tanto sueño que no podía abrir (*se) los ojos	He was so sleepy that he couldn't open his eyes
71	Se levantó la pierna porque la tenía dormida	He lifted his leg because it was numb
72	Amaneció con una infección en los ojos y {se/* \emptyset } los tuvo que abrir con los dedos	He woke up with an infection in his eyes and he had to open them with his fingers
73	{ \emptyset_x /se; _i } \emptyset_i sirvió _i una copa _i	He _i served a drink _i (to someone _x /himself _i)
74	{ \emptyset_x /te; _i } \emptyset_i preparaste un café _i	You _i prepared a coffee _i {for someone _x /yourself _i }
75	Mariana se quitó la falda	Mariana took off {her/the} skirt Removed her skirt/Removed the skirt from herself
76	Rasgó las vestiduras del auto	He ripped the car's seats
77	Le rasgó las vestiduras al auto	He ripped the car it's seats
78	Puso las luces en el árbol	He put the lights on the tree
79	Le puso las luces al arbol	He put the tree some lights
80	Le cambiaron las llantas al coche	They changed the car's tires
81	[FR] Je {*lui/en} ai oublié le nom	I have forgotten his/its name
82	[IT] {*gli/ne} ho dimenticato il nome	I have forgotten his/its name
83	[SP] Le _i puse el mantel [a la mesa] _i	I put the tablecloth on the table (Demonte 1995:12)
84	Valeria le _i miró las llantas [al auto] _i	Valeria looked at the car's tires (Cuervo 2003:78)
85	Ya le _{SG} dieron a los niños _{PL} su pastel	They already gave them their cake

Affectedness can be extended to inanimates as whole-part construals (76-80). In languages with adverbial clitics, non-affecting verbs can only appear with 'genitive' *ne/en* (81-82, Belletti & Rizzi 1981; for French, Kayne 1977:§2.15; Vergnaud & Zubizarreta 1992:§1). Spanish (Picallo & Rigau 1999; Sánchez López 2007) 're-uses' *le* (83-84). In Mexican

Spanish, *le* need not show number agreement with its complement. This appears to be an incipient [-SPEC] form, also found acting as a sort of locative (Maldonado 2002b). Butt & Benjamin (1994:141) discuss uses of Spanish *le* similar to those Italian *ci*_{IMP}.

In addition, French/Catalan/Italian may use locatives as inanimate/unspecified datives to highlight lack of affectedness, which in Spanish is expressed through not doubling the clitic (see below).

3.2.5 Clitic Doubling

Affectedness is further highlighted in Spanish³⁴ by ‘dative-doubling’ which is so common as to be considered almost ‘obligatory’, however, complements may occur without clitics, especially in formal/written discourse. As DAT-ACC relationships become looser, the possibility of omission increases e.g. (86) where transfer is abstract since the recipient is a replicate mass. Introducing *les* implies that speaker and audience made eye contact. Where such contact is required, omission is unacceptable (87-88). At the discourse level, dative-doubling allows speakers to validate events: in (89) without *le*, the subjects simply agree on their support; with *le*, they actually expressed it to the candidate and the speaker validates such actions from his own experience or that of an unquestionable source. Even in cases of real transfer, it remains possible for conceptualizers to refrain from validating (thereby establishing distance from) the event by clitic omission, as observed in newspaper headlines (90) and formal/reported speech which tend towards omission even for well defined participants (91, Delbecque & Lamiroy 1996).

34 Standard French/Italian does not accept dative doubling. Romanian doubles DAT_[+E], but not DAT_[-E].

Table 30		
86	{les/Ø} pidió a los manifestantes que...	He asked the protesters to...
87	{le/'Ø} dio un beso a Adrián	She gave a kiss to Adrian
88	{le/*Ø} quitó las monedas de la mano	He grabbed the coins from his hand
89	(le) Ø _i manifestaron su apoyo al aspirante	They showed their support to/for candidate
90	Ø _i Ø _i dieron el Nobel _i a García Márquez _i	They gave García Márquez the Nobel price
91	Ø _i Ø _i atribuyen la paternidad _i a Juan _i	They attribute paternity to John

3.2.6 Conclusions for the Lower Clitic-Field

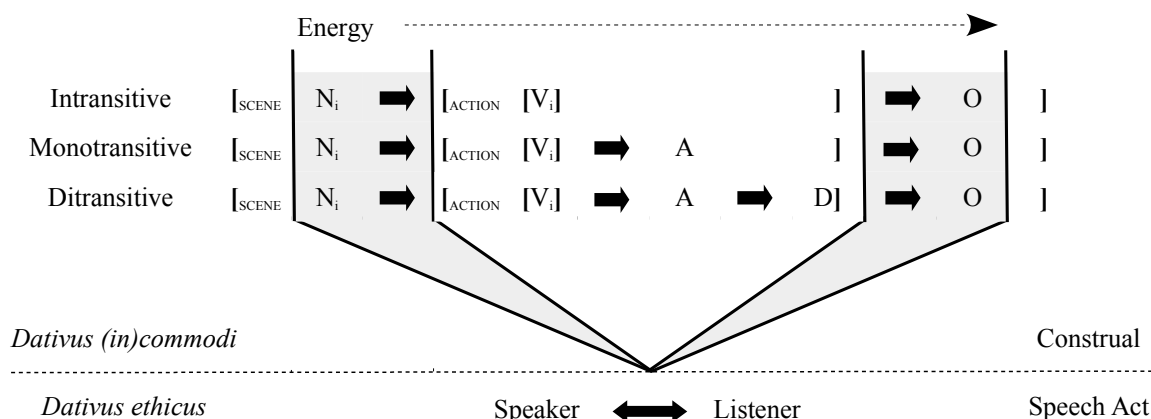
There is a syntactic requirement (modulo generic cases discussed in §2.2.3) that arguments should be filled, but most uses of DAT clitics do not come under this heading. Even with recipient/source argumental datives, syntax requires the presence of an argument, but if it is expressed as a complement, it may or may not appear as a clitic as well. Far from a clear cut analysis based on simple ‘possession’ with specifiable ‘rules’, DAT (like OBL) is vague. Its function is merely to indicate the presence of an affectee within the event. The relationship between this affectee and the effectee has to be inferred.

If each sentence is read in context, no rules/stipulations are required; in fact, they only lead to error, because they seek to make a requirement of what is merely the default reading/situation and, thereby, incorrectly reduce the range of meanings actually found in real usage.

3.3 Upper Clitic-Field

Whilst the lower-field expresses relationships between event participants, the upper-field introduces participants external to the inner action: the *effector* imparting energy into the event, and additional participants who remain out of the spotlight but incorporated into the wider *scene* to depict their evaluation of, or affectedness by, *the action* taking place under the

spotlight.³⁵ Without NOM/OBL, the spotlight is upon the action e.g. arrival of ACC with DAT, the subject is present within the verb, but not treated as part of the focus; with NOM/OBL, the spotlight expands to include the subject's relationship to that action e.g. NOM's giving ACC to DAT,³⁶ and/or that of third parties experiencing/evaluating that event.



Unlike DAT, OBL cannot be subject coreferent nor, being out-of-the-spotlight, coreferent with participants. Many authors divide OBL between ‘sympathetic’ vs. ‘settings’ datives, although the dividing line varies between authors. In our model, the division is represented by $[\pm E]$.

3.3.1 Sympathetic

Sympathetic datives may be omitted without major change in sentence meaning, sometimes described as “superfluous” (Bello & Cuervo 1960) or “procedural” (*sensu* Sperber & Wilson 1988) i.e. not contributing to sentential truth conditions, but rather expressing attitudes. Whilst the *affectee* within the event *le...a Valeria* is an object-dative (92), *me* introduces a non-participant (not necessarily present) who intellectually evaluates the event from their perspective. §3.2.5 showed dative-doubling in the lower field as an evaluation of the speaker’s understanding of propositional veracity; here, it relates to the event’s impact. They

³⁵ §4.7.1 for the NOM~OBL distinction.

³⁶ i.e. their role is heightened in listener awareness, as seen in the effects of nominative SE_{NOM}/SE_{ANT} (§4.7).

differ from *dativus ethicus* which are external to context, referencing participants within the speech-act, temporarily bringing the conversation out of discourse and into speech-act here-and-now (§2.1.2). Sympathetic datives reference non-participants within the construal, not as interlocutors, but as their projections within the construal i.e. their on-stage role. Similar usages are found in all Romance languages e.g. (93-98).

Table 31

92	[SP]	Me le arruinó _i la fiesta _j a Valeria _k	He ruined V's party on me
93	[CA]	No te m' enfadis	Don't get angry on me
94	[FR]	Jean lui a mangé tout le fromage	J. ate all the cheese on him/her
95	[RO]	Vor să mi ti omoare	They want to kill you on me
96	[IT]	Juan me le ha rovinato la vita (a quella ragazza)	J. has ruined her (that girl's) life on me
97		Mi ti vogliono uccidere	They want to kill you on me
98		Jean gli/le ha mangiato tutto il formaggio	J. ate all the cheese on him/her

Strozer (1976) opines that sympathetic datives require presence of object-datives (100); without them (101), *me* must be read as DAT. This description is too strong. With both datives present, position determines each role. With only one, vagueness tends to be resolved with DAT readings. This follows from evaluation order from inside outwards (((((V)A)D)O)N). For ditransitive verbs, ACC then DAT must be filled, for OBL to be recognised as such. However, for monotransitives, it is possible to 'skip over' optional DAT, and read single datives as OBL, where possession is contextually inappropriate (102), or possessive adjectives 'fill the gap' (103~104). Even with ditransitive verbs, context alone may be sufficient for OBL readings, (105-106).

Table 32

	O	D	A			
99	me	le _k	Ø _j	arruinó _i la fiesta _j a Valeria _k	He ruined V.'s party on me	Spanish
100	me _i	le _k		comió _i la hamburguesa _j (a V) _k	He ate V.'s hamburger on me	
101		me _i		comió _i la hamburguesa _j	He at my hamburger/*on me	
102	me _i		Ø _j	detuvieron [a los rateros _j]	They stopped the thieves _j for me	
103		me/te/le		arregló _i la ventana _j	He fixed (my/your/his) window. DAT=owner	
104	me/te/le	Ø _k		arregló _i su _{i/k} ventana _j	He fixed his window for/on me. OBL≠owner	
105		le _i	Ø _j	alquila la casa _j a Pablo _i	She rents the house;	from/to/of/for Pablo _j
106	le _i					on Pablo _i (against his _j wishes)

For French, Herschensohn (1992, *i.a.*) argues that sympathetic datives must be linked (usually possessively) to ACC. Authier & Reed (1992) present a different interpretation. Such datives are regularly found with transitives without any relationship between dative and verbal object (107), but not unergatives (108) unless used transitively (109). Nor is ACC required. Subcategorized oblique objects (110), and VP-internal adjuncts (e.g. locative/manner PPs, 111) also license such datives. VP-external adjuncts denoting cause/time, or simple adverbs do not (112). Nor is ACC sufficient. Idioms (113, Rouveret & Vergnaud 1980:170) are unacceptable, but become so when additional place complements render the event specific (114). Nor are circumstantial adjuncts adequate in themselves. They must be salient, highlighting the process' pertinence to the sympathetic referent. In (115), *dansé* is habitual having no consequence upon the clitic's referent without further specification (116).

Table 33

107	Je vais te lui écrire une lettre	I'm going to [write a letter to him] for you	French
108	*Paul lui a bu	Paul [drank] on him	
109	Paul lui a bu trois pastis	Paul [drank three pastis] on him	
110	Il lui a parlé à sa fille	He [spoke to his daughter] for him	
111	Alfred lui a roté devant les invités	Alfred [burped in front of the guests] on him	
112	*Alfred lui a roté pour choquer ses invités	Alfred [burped on him] to shock his guests	
113	*Il lui a cassé la croûte	He [had a bite to eat on him/her]	
114	Il lui a cassé la croûte sur ses beaux coussins de cuir	...[bite to eat on his nice leather cushions] on him	
115	*Il t'a dansé	She [danced] for you	
116	Il te <l'> a dansé <un très beau tango/ça>	She [danced it/a very beautiful tango] for you	

Contra Herschensohn (1992), Authier & Reed (1992) argue that sympathetic datives consistently refer to individuals understood as being concerned by the event as a whole and as such, do not form θ -chains with empty categories within VP. As with lower-field ‘possessives’, *dativus (in)commodi* appear to be ‘applied’ arguments. As such, achievement of this reading is dependent upon the nature of the event to which it is applied: ACC must be specific to be possessed by DAT, VP must be ‘specific’ for OBL to ‘possess’ it. It is from pertinence that ‘possession’ may be inferred as contextually appropriate. Pertinence is more difficult to show in, and hence sympathetic datives are rarely found with, intransitives which represent internal states. OBL may occur with such verbs when it bares a clearly evaluative character i.e. settings datives (§3.3.2). As with object-datives, there are no ‘rules’ or structural implications, merely appropriateness.

Table 34

117	[SP]	Mi bebé me lloró toda la noche	El bebé del vecino *me lloró toda la noche
118	[IT]	Il mio bambino mi ha pianto tutta la notte	Il bambino del vicino mi ha pianto tutta la notte
	[EN]	My baby cried on me every night	The neighbour’s baby cried on me every night

Roberge & Troberg (2009) provide similar cases in Italian/Portuguese, whilst noting that Romanian/Spanish are not effected by such restrictions. We interpret this as a language-specific phenomenon overlaid upon the simpler cross-linguistic (i.e. structural) pattern. Indeed, Shibatani (1994:464) who considers use of sympathetic datives to be motivated by ‘proximity’ to, and ‘relevance’ of, the event, shows that there are cultural differences of acceptability even for the same sentence across languages e.g. Spanish vs. Italian (117-118, Shibatani 1994:472-473).

3.3.2 Settings

Settings datives relate to events, defining mental-locations where the event has significance.

The *gustar* class of verbs sometimes termed “impersonals” (RAE 1973:§3.13.4), or “inverse verbs” (Delbecque & Lamiroy 1996; Vázquez 1995) depict human dispositions, selecting ‘dative’ arguments with ‘experiencer’ θ -role (Belletti & Rizzi 1988). The non-active subject (hence, S_L position) is source of an on-going emotional state (intransitive VP) within OBL’s dominion (119) i.e. OBL ‘possesses’ the affects of the event. More rarely, these verbs occur without OBL (120-124, Sánchez López 2007) inducing generic (\emptyset_{OBL}) readings. Equally, subjects may be omitted (122).

Table 35

119	Me gustan los libros	I like the books	Spanish
120	\emptyset_x gustan los libros	Books are well-liked (in general)	
121	En estos actos siempre \emptyset_x duele la cabeza	In these kinds of events, one always has a headache	
122	¿Quieres este helado? _i ? No, no me gusta \emptyset_i	Do you want this ice cream? No, I don’t like it	
123	Le falta la sal	Salt is missing on/in it	
124	\emptyset_x falta la sal	The salt is missing (i.e. on everyone)	
125	Adoro los libros	I adore books	
126	María se gusta mucho (a sí misma)	M. likes herself a lot	
127	(Le) es difícil aceptarlo	It is difficult (for him) to accept it	
128	Me gusta {el pelo/mi pelo/tu pelo}	I like my/your hair	
129	Me duele {la cabeza/en la cabeza/*mi cabeza}	I have a headache	

Such constructions are often treated as equivalents of active constructions i.e. (125)≈(119), but (125) is transitive, whilst (119) is stative. Just as Old English *like* previously had different argument structures (*him like oysters* vs. *he likes oysters*, Jespersen 1924:160), *gustar* was transitive in Old Spanish, coexisted with patterns with prepositional objects (XVI^c, continued in European Portuguese), and became expressed with OBL by the XVIII^c: a “semantic change...from a tasting agent to a satisfied experiencer” (Whitley 1998:138). Transitive uses still survive (126), which represent different construals (closer to 125). Van Valin & La Polla

(1997:154) liken the distinction to English *own=have* vs. *belong to* (predicate of state). Often treated (unnecessarily, from our perspective) as a separate/special group, these verbs simply shifted from predominantly transitive to predominantly intransitive uses, joining many verbs already following this construction i.e. (119) is no different from (127).

Any ‘possession’ is inferred as shown by the contrasts between physical (*doler*) vs. psychological (*gustar*) experiencer verbs. In (128), *el pelo* is generally inferred as belonging to the experiencer, but other people’s hair may be the subject (*tu pelo*), so that some circumstances may require specification (*mi pelo*). In such cases, no dissonance is caused by its inclusion. In (129), however, the pain of other heads cannot be experienced and so the possessive adjective is questionable outside of the mortuary scenario of §3.2.1. In both cases, dative=experiencer; possession by the subject is inferred or, where appropriate, denied (*tu pelo*). Mexican Spanish (Maldonado 2002b) follows the reverse logic. (129) with a possessive adjective is commonplace. Confusion would require very particular context in (128), and is impossible in local-person (129). The possessive adjective is ‘superfluous’ but not considered misleading (by speech community convention) and is, therefore, optionally available to add emphasis, or invoke empathy e.g. (129, addressed to a loved-one, not medical professionals).

Some analysts treat OBL in these as ‘dative subjects’. Campos (1999:1,560) raises coreferentiality tests with temporal infinitival constructions, where the datives of these verbs control the infinitive’s subject (130) as subjects may in dynamic situations (131), however, the putative subject is unable to control the adjective in (132). All that can be gained from such tests is that NOM *and* OBL (both IP participants) are structurally ‘high’.³⁷

37 Comrie (1981:53-6)’s control continuum places experiencers closer to agents, and separated from patients.

Table 36

130	A Luci _i le gustaba Ronny _j antes de e _{i,*j} conocer a Otto	L. liked R. before meeting O.
131	A Ronny _j le escribía Lucy _i antes de e _{i,*j} conocer a Otto	L. used to write to R. before meeting O.
132	A María _i Juan _j le desagrada borracho/*borracha _i	M. dislikes J. drunk

3.3.3 State, *not* Place

While verbs like *sentir* (133, and those of the previous section) result in stative predicates, achievement verbs (134) and anticausatives (135) produce COS predicates emphasizing initiation of a new state. Settings datives indicate union of the referent with that state, and in this sense, personal OBL ‘possess’ the affects of the event. Languages with ‘adverbial’ clitics (Chapter 5), locative *ci/y* is treated as the state with which the event is associated, and ablative *ne/en* as the state left behind in order to achieve the new state. Alternatively, such non-personal clitics may anaphorically reference individuated places. Personal OBL cannot, and cannot be used as destinations (136) or sources (137). Similar looking uses are allowed where they indicate an experiencer of the event’s affects (138-139); even (140) is acceptable for some speakers/dialects.

Table 37

133	Le sienta bien el vestido	The dress sits well on her	Spanish
134	Le entraron ganas de llorar	A crying feeling entered him	
135	Se me murió	He died on me	
136	*M’ha venido?	*He came to me	
137	*Le fue	*He went from him	
138	Le fue bien en Buenos Aires	It went well for her in BA	
139	Al perfume se le fue el aroma	The lotion let go its odour	
140	%Ya le camina	She is already walking for him	

In (141-142), the adverb *encima* indicates ‘(from) above’. An agent/cause (hence, in S_H) achieves an internal change-of-state of (dis)position (hence, anticausative marker SE_{ANT}, see §4.7.3). The affects of the event can be experienced by third-parties, for which possession

may be inferred (141), or not (142-143). In (144), *encima de mí* means ‘on top of me’. Some analysts link (143) with (144) by a process of extraction of the pronoun from the adverbial phrase. For French (149-150), Kayne (1975:158) suggests that such datives should be considered realizations of obligatorily affected internal ‘locative’ arguments.³⁸ Verbs like *pasar* when denoting achievements allow internal (145) or external (146) realization of these arguments, but activity movement verbs, like *caminar* (147) which do not imply change in locative relation between the involved arguments, nor movement verbs involving a change of locative relation when in stative constructions (148), do not. But this does not amount to extraction, merely that such verbs have such an argument slot available.

These are simple manner adverbs which support an optional adverb-internal argument. OBL indicates a participant in union with the event. As indicated in the translations (149-150), the implication of motion towards/away from that participant is inferred; there is no need for the subjects of (149-150) to come into contact with *lui* at all, whilst the adverbs retain their meanings ‘downwards’ and ‘inwards’. (143~144) and (145~146) are separate construals expressed in distinct syntax. The fact that their meanings can overlap (or be interpreted to do so) does not warrant extension of theory to include extraction from doubly subordinated clauses. These extended adverbial phrases are clarifying functions much like the reflexive emphatics discussed in §3.4.4. The OBL in these examples, therefore, remains an experiencer, *not* part of a split locative expression.

38 See below for Italian examples.

Table 38

141	El mundo se le vino encima	His world came (tumbling) down	Spanish
142	La noche se nos echó encima	Night fell (suddenly) on us	
143	El gato se me sentó [ADV encima	The cat [sat down] on me	
144	El gato se Ø sentó [ADV encima [PP de [NP mí]]]	The cat [seated itself] on me	
145	Le pasó [ADV por delante]	He passed in front of him	
146	Ø pasó [ADV por delante [PP de [NP él]]]		
147	*Le camina delante	He walks in front of him	
148	*Le está sentado encima	He is seated on him	
149	On lui tombe dessus	They are falling on top of her	French
		They are falling down on her (against her best intentions)	
150	Le couteau lui entre dedans	The knife goes into him/her	
		The knife goes inwards on her (e.g. into her best settee)	

Anticausative uses are common (151-152), where SE_{ANT} indicates culmination of a prior state and ingress into a new state, driving expectations that someone may be *affected* by such (often abrupt) changes-of-state. Location may be profiled (151), or not (152). Again, experiencer is quite distinct from any attendant locative adjuncts e.g. (153), where *de las manos* describes subject trajectory. Note that any ‘possession’ is inferred as shown by (154). In the absence of SE_{ANT} , the effect depends on verbal semantics, ranging from unacceptability (155-156) to reading as a sympathetic dative (157-158, 159-160).

Italian/French lack equivalents of (151-152). This is not, however, a clitic~clitic restriction. Italian/French allow adverbial (i.e. impersonal) clitics with SE_{ANT} (161) where Spanish (lacking adverbial clitics) leaves the space empty, but with strong implication of source (§4.7.3). As noted by Schäfer (2008), personal OBL cannot pronominalize in these languages, with marked (162) or unmarked (163-164) anticausatives, whilst nonetheless carrying the same inferences including extended ‘unintentional causer’ readings (§3.3.6). Hence the lack of $SE_{ANT}+OBL$ in these cases is part of a wider language-specific restriction, rather than a local clitic restriction.

Table 39

151	Se me murió en las manos	He died in my hands	Spanish
152	Gonzalo se me volvió loco	Gonzalo went crazy on me	
153	La pelota se le cayó de las manos	The ball fell from the hands on him (→his hands)	
154	La pelota se le cayó de las manos de Juan	The ball fell from Juan's hands on him	
155	Se me murió	He died on me	
156	*Me murió		
157	Se me cayó	It fell from me	
158	Me cayó	It fell {on/*from} me	
159	*Le fue		
160	Se le fue	He went away on him	
161	Se ne va	He sets off	Italian
162	A Franco; si Ø _i rompe il vaso (per errore)	The vase broke {on/because of} F _i	
163	A Franco; Ø _i è bollito fuori il latte (per errore)	The milk boiled over {on/because of} F _i	
164	A Franco; Ø _i sono appassite le piante (per errore)	The plants wilted {on/because of} F _i	
165	Le siedo vicino a Giulia	I'll sit near to Giulia	
166	Ci/*le siedo vicino alla porta	I'll sit near to the door	
167	*Mi le siedo vicino a Giulia	I'll sit myself near Giulia	
168	Mi ci siedo vicino alla porta	I'll sit myself near the door	
169	Se ci siede vicino alla porta	He'll sit himself near to it	

Like Spanish, French/Italian allows participants with intransitives accompanied by manner adverbs, expressing a third-party externally (subordinated to the adverbial phrase) or internally as OBL (examples from Pescarini 2015). The participant may be animate (*le*, 165) or inanimate (*ci_{IMP}*, 166). As already indicated, when SE_{ANT} is present, the combination *me+le* (167) are not available, but *ci_{IMP}* is (168). Given the arrangement in (168), (169) should be possible producing a SE+*ci* sequence. Comparable forms are found in languages using *y* rather than *ci* (e.g. Aragonese, §6.6), however, we have never seen (168) or (169) in use.

3.3.4 Possession

(170) introduces an affected participant. When the *eyes* are known to be separate from that participant, benefactive/malefactive readings (OBL) are inferred (170a). Possessive readings are expected when there is a part-whole relationship between affectee and object (170b, determined by discourse and/or world knowledge). In traditional terms, this is possible

because DAT c-commands ACC. Possessive readings are still possible if context forces a benefactive reading (170c), because OBL being higher in the syntax tree also c-commands the event as a whole. In (171), the same relationship holds between OBL and the ‘patient’ subject in S_L . (170-171) should be compared with (172) where the logical subject, having been removed to an adjunct clause, does not c-command the logical object (grammatical subject) and possessive readings are unavailable. Structure defines affectee~effectee relationships. The function of structure is *not* to define the nature of the affectedness, merely its existence. The hearer infers whatever is appropriate to the situation in terms of possession/ownership.

Table 40

170	Maria me _i cerró los ojos _j	a. Maria closed the eyes _j on/for me _i (e.g. eyes of a doll) b. Maria closed my _i eyes _j c. Maria closed my _i eyes _j for me _i (I am unable to do so)
171	Se te _i ha arrugado la piel _j	The skin _j has wrinkled on you _i (The one on the table) Your _i skin _j has wrinkled Your _i skin _j has wrinkled on you _i (your skin and that affects you)
172	La cabeza _j fue levantada (por Juan _i)	The/*his _i head _j was lifted (by Juan _i)
173	Pablo le puso azúcar al mate	Pablo put sugar in the tea
174	A la mesa se le rompieron dos patas	Two legs of the table broke

Both ‘datives’ can be found with inanimate entities. For DAT (173, Cuervo 2003), the meaning conveyed is that the non-human dative has/possesses the entity expressed by ACC after the event has taken place. For OBL, (174) expresses that the inanimate entity has/possesses the new resultant state. This reading is only possible with inanimate datives when a relation of possession can be implied as (the only possible) source of affectedness (McIntyre 2006).

Table 41			
175	Mi scrivi questa lettera?	=Scrivi questa lettera... al posto mio?	Italian
176		...per me?	
177		...a me?	
178	Mi hanno ucciso la figlia	They killed the daughter on me (I was responsible for her)	
179	Mi hanno ucciso mia figlia	They killed my daughter on me	
180	Gli è morta la mamma	The mother died on him	
		His mother died	
181	A Gabi le llegaron dos cartas	There arrived two letters on Gabi	Spanish
182		There arrived two letters for Gabi (implied receipt)	
183	Je lui ai lavé la/sa voiture	I washed the car on him	French
184		I washed his car	

When only one dative form is present, more than one interpretation is often possible. As Simone (1993:97) notes for Italian, this can lead to three way ambiguity of surface-forms (175). This is particularly common with inferences of possession (178). In such circumstances, possessive adjectives may be used to clarify the situation (179), even where such specificity is usually avoided (O'Connor 2007). Hoekstra (1995:127) makes similar comments for French: in (183), the possessive adjective in *lui...sa* forces a reading of *lui*_{OBL}, whilst *lui...la* is read as possessive. In intransitives, the clitic must be OBL (since intransitives lack D/A structures), but possession (181) and even reception may still be implied (182). Neither are inherent in the structure, merely inferred.

3.3.5 Restrictions

In Standard Spanish, settings (185-186), but not sympathetic (187), datives may be doubled by PPs (Strozer 1976; Jaeggli 1982). Franco & Huidobro (2008) associate this with argument status: settings datives are arguments, sympathetic datives are applied (i.e. the same relationship between goal DAT and static/possessive DAT). However, there is dialect/idiolect variation in acceptability (Roldán 1972:30-31), such that matters cannot be so direct.

There are no person restrictions on settings datives (188). For Standard Spanish, Bello & Cuervo (1960) consider (repeated Strozer 1976, *i.a.*) that sympathetic datives are limited to 1-person, but other persons do occur (189, Argentinian, García 1975). For Mexican Spanish, (Maldonado 1992, 1999) illustrates a 1»2»3 subjectivity hierarchy (190). The event and its effects are linked to a conceptualizer, which is normally the same as the speaker. When speaker empathizes with hearer, sympathetic datives may take 2-person. Even more rarely, this may be extended to 3-person. This is not a person restriction in terms of clash of clitics or syntactic property, but rather a naturally skewed distribution based on discourse behaviour. Humans are most interested in what they think/feel themselves, possibly what their interlocutor thinks, but rarely the emotions of outsiders.

Table 42

185	Se nos murió a nosotros	S/he died on us	Spanish
186	Se le quedó dormido a su madre	He went to sleep on his mother	
187	Me _j le _i arruinaron la vida a mi hija _i *a mí _j	They ruined my daughter's life on me	
188	Se me/te/le(s) rompió	It broke on him/her/you/me/them	
189	Te le arruinaron la vida a tu hija		
190	Les/me/te/nos galardonaron al presidente	They gave an award to the president on them/me/you/us	
191	Te le han dado un premio a tu hija	They have given a prize to your daughter on/for you	
192	Me castigaron al niño	Peninsular: They punished my son Latin American: They punished the kid on me	
193	Me le pusieron un cuatro al niño	They flunked my son (gave a fourth to the kid on me)	
194	No le duerme	He doesn't sleep for her (Cuervo 2003)	
195	Juanita ya le camina	Juanita can already walk on him/her	
196	El niño le estudia bien a Maria para los exámenes	The boy studies hard for his exams for Maria	
197	Me le _i dieron un helado al niño _i	They gave the kid an ice-cream on me	
198	¿Te lo llamo al doctor?	Should I call the doctor for you?	
199	Il mio bambino non mi dorme	My baby won't sleep for me	Italian
200	*Questo bambino non ti/gli dorme proprio	This baby won't sleep for you/her	

OBL invokes interlocutor empathy generally driving negative inferences, but positive evaluations are also possible (190-191). In Ibero-Spanish, positive readings (other than benefactives) are almost consistently rejected under elicitation, yet commonly heard in spoken

informal situations (Maldonado 2002a). In most Latin American dialects, (192-193) are acceptable. In more conservative dialects, (192) is only acceptable when read as possessive DAT, whilst external participant *me* (193) or the alternative reading of (192) are banned. Similarly, conservative dialects tend to employ only 1-person, while use of 2-person is more frequent in less restrictive ones. Lack of Ibero-Spanish sympathetic *le*_{OBL} appears to be quite robust. We consider absolute clitic availability to reflect each dialect's clitic lexicon i.e. only some speakers possess sympathetic *te*_{OBL}, very few *le*_{OBL}. Chapter 7 shows the crucial nature of the availability of clitics when considering putative PCC-breaches.

Italian unergatives are commonly used with OBL (199), where *mi* is not experiencing the child's lack of sleep, but evaluating the effect of such (repeated) behaviour, as in Spanish (194). The standard language is limited to 1.SG, but some dialect/idiolects do accept (200, Roberge & Troberg 2009). Like Spanish, Italian displays dialect/idiolect-dependant *mi~mi/ti~mi/ti/gli*. French follows a similar pattern.

3.3.6 Inferences of Causation

In Spanish,³⁹ neither marked (201) or unmarked (202) anticausatives, nor non-alternating unaccusatives (203) license *by*-phrases introducing external arguments, but all three license an extended range of readings for OBL.⁴⁰ This appears to hold across all languages (Alexiadou *et al.* 2006a, 2006b).

Schäfer (2008:69) claims that sentences like (203) contain “unintended Causer” readings. Cuervo (2003) merely claims “unintentional responsibility” (Cuervo 2003:187). Fernández

³⁹ Data from Fernández Soriano (1999).

⁴⁰ As noted above, Italian/French are unable to show all of these clitic patterns, but the same readings are available with complements.

Soriano (1999:134) reads them as simple benefactive/malefactive. The term ‘unintentional causer’, although frequent in the literature, does not capture the range of meanings found across languages. In (204), the girl may be unintentional causer (204a), involuntary/indirect facilitator (204b) or unexpected causer (204c), depending on contextual/pragmatic factors. Canonical transitive subjects may act unintentionally or accidentally as suggested by adverbs in (205), but the other readings do not obtain. Anticausatives, however, which imply spontaneous action, may take such readings, unless a cause is indicated (206/207).

With transitives, a cause(r) is present (taking nominative), such that no other cause(r) can be introduced; OBL may only take experiencer/evaluator readings. In externally-caused anticausatives/unaccusatives, the nominative represents an agent (*in sensu* Higginbotham (1997)’s teleological capabilities), but the semantic role of cause(r) is empty. Only in these cases, may the role be inferred (or transferred to) OBL (McIntyre 2006:204).

Internally-caused COS verbs, by definition, already have a cause(r) and, therefore, implication of another necessarily external cause(r) is impossible. Extended readings require a ‘possessive’ relationship to be inferred, from which an element of ‘responsibility’ for the COS might be understood. Note that (208-209) would be unacceptable if context determined the nose/double-chin belonged to someone else. Similarly, cases where bare NPs are acceptable are those where possession is inferable (210). OBL is read as an entity capable of creating the environment in which the internally-caused COS takes place hence the impossibility of (211), but acceptability of (210). Many internally-caused COS verbs do not normally admit OBL e.g. *oscurecer* (212), but do so as marked anticausatives (213). The effect of SE-marking is to indicate that such ‘responsible’ actors might be inferred.

Table 43

201	A Juan, se le rompieron las gafas	The glasses broke {affecting/because of} J. _i
202	A Juan, le hirvió la leche	The milk boiled over {affecting/because of} J. _i
203	A Juan, le florecen los árboles	The trees bloom {benefiting/because of} J. _i (=good gardener)
204	A la niña, se le , abrieron las puertas	a. The girl accidentally caused the doors to open b. The girl let the doors open c. The girl managed to open the doors, unexpectedly
205	La niña abrió la puerta sin querer (al apoyarse)	The girl opened the door accidentally (by leaning on it)
206	Al chef, le , quemaron la comida: fue el pinche.	The food got burned on/affecting the chef: it was the scullion
207	Al chef, se le , quemó la comida: #fue el pinche	The food got burned because of the chef: #it was the scullion
208	A Pinochio, parecía crecer- le , { ^l la nariz/#el pollo}	The { ^l nose/#chicken} appeared to grow on Pinocchio (His nose, because of lying)
209	A María, parecía engordar- le , la papada/#el pollo}	The double-chin appeared to grow on M. (Her double-chin, because of over-eating)
210	A Juan, le , crecen flores en el pelo	Flowers grew in John's hair
211	A Juan, le , brotan champiñones *(debajo del brazo)	Mushrooms grew under John's arms.
212	#A Juan, le , oscureció el día	#To Juan darkened the day
213	A Juan, se le , oscureció la plata	The silver darkened on J.
	...porque le echó un producto corrosivo	...because he applied a corrosive product
Caer Unaccusative 'Fall'		
214	Me cayó un plato (encima)	A plate fell (from above) on me } to my disadvantage, Lightning struck/fell on me } <i>not</i> in my direction
215	Me cayó un rayo	
Caer(se) Anticausative 'Drop'		
216	# Se me cayó un rayo	#The bolt of lightning dropped (on me) #I let a bolt of lightning drop (accidentally) #The bolt of lightning dropped (despite my intentions)
217	Se me cayó el plato (de las manos: source)	The plate dropped (on me) I let the plate drop The plate dropped (despite my intentions)
218	A la olla, se le , cayó el asa	The pot's handle dropped off
219	A Juan, se le , cayó el libro	J. let the book drop J. accidentally dropped the book The book fell on/affecting John
220	A Juan, se le , cayó el pelo	John's hair fell out (affecting him)
221	A la muñeca, se le , cayó el pelo	The doll's hair fell out
222	Al cepillo, se le , han caído los pelos	The hair dropped from the brush

Caer may optionally appear with SE. When items fall naturally, SE is unavailable, as are any extended readings (216), only affectee readings are available (214,215). With SE, the anticausative introduces the possibility of other readings, including an ablative quality (218), where prior possession/proximity is inferable (217,218). See §3.3.3 for arguments against these being truly 'locative'.

Availability of extended readings is determined by structure. Appropriateness of inference is controlled by the nature of the participants. The restriction to humans is because only humans can be intentional. The restriction to certain objects is due to world knowledge of what such an intentional causer is capable of intending. Such inferences depend on the perceived (i.e. the interlocutor inferred) relationship between the entities involved and world knowledge of their capabilities. (219-222) have the same structure, but different sets of readings are available in each. Its syntactic presence as OBL indicates a participant which is significance for the event. The hearer determines that significance from context. That OBL must be human in these circumstances merely shows that the participant must be capable of the property which is attributed to it.

3.4 Separating Fields

There are several phenomena which appear to breach the OBL~DAT divide. Ignoring the OBL~DAT distinctions leads to the definition of putative problems which require complex approaches to solve. This section explores a number of areas where a clear understanding of the difference can in fact simplify our understanding of this area of investigation.

3.4.1 Absence of OBL_[+R]

Direct-objects can be passivized across ditransitive (223), or monotransitive (i.e. applied possessive) DAT (224), but not over OBL (225, Demonte 1994). Passives do not license DAT: *le* in (226-227) is OBL, hence SE-reflexive (228) and SE-reciprocal (229) are unavailable. Nor is DAT available with copulas i.e. intransitives lacking DAT/ACC structure, only NOM/OBL. (230) is marginally acceptable as an OBL affectee, but not DAT recipient. Again, since OBL has no reflexive, reciprocal (231) or reflexive (232) readings are impossible. Even where OBL is expected, it cannot be reflexive (233). Similarly, in (234~235, Rizzi 1986)

where the ‘dative’ clitic as DAT would breach the PCC, and raising verbs (236~237, Burzio 1986).

Table 44

223	El premio Nobel le fue concedido a Cela el año pasado	Spanish
224	La pierna le fue vendada a Pedro cuidadosamente por el doctor	
225	*Mi niño me ha sido suspendido otra vez por ese profesor	
226	El professor le ha sido presentado (a M.)	The professor was introduced to M. Spanish
227	J. y M. le han sido presentados (a K.)	J. and M. was introduced to K.
228	El profesor (*se) ha sido presentado (a sí mismo)	The professor was introduced to himself
229	J. y M. (*se) han sido presentados (el uno al otro)	J. and M. have been introduced to each other
230	*?J. le es cruel a su vecino	J. is cruel {*/to/'on} his neighbour
231	*J. y su vecino se son crueles	J. and his neighbour are cruel to each other
232	*J. se es cruel (a sí mismo)	J. is cruel to himself
233	(A J.) le/*se es fácil resolver estos problemas	It's easy for J./*himself to solve these problems
234	G. <gli> è stato affidato <a lui>	G. was entrusted to him Italian
235	G. <*si> è stato affidato <a se stesso>	G. was entrusted to himself
236	Jean leur semble intelligent	Jean seems intelligent to them French
237	*Jean se semble intelligent	Jean seems intelligent to himself
238	I. și M. și-au fost prezentați	J. and M. were presented to each other _i Romanian

Alone of all the Romance languages, Romanian possesses personal locative clitics (i.e. DAT_[+E]) which are available in passives, like non-personal locatives (238, Dobrovie-Sorin 2006:132), allowing Romanian to apparently breach the PCC (§7.4.4).

3.4.2 Laísta Dialects

Some languages show differences in form between OBL and DAT. In Standard Spanish, both OBL and DAT 3-person is represented by *le(s)*, regardless of gender (241-244). In *laísta* dialects (Romero 1997, 2001), *la(s)* represents both ACC (240) and DAT (239) feminine referents, but OBL retains *le(s)* (242-244). Note that in (242), *le* cannot be DAT since this position is filled by *a casa*, and benefaction rather than reception is indicated. As expected, *la*_{DAT} cannot appear with passives (243, Gutiérrez Ordóñez 1999:1870), or unaccusatives

(244). Contra Romero (2012), this is not evidence that la_{DAT} is really accusative, but simply shows that *laísta* dialects have clitic paradigms with $le_{\text{OBL}} \sim la_{\text{DAT}} \sim la_{\text{ACC}}$ instead of standard $le_{\text{OBL}} \sim le_{\text{DAT}} \sim la_{\text{ACC}}$. This underscores the fact that we must rely on functionality (and when two datives are present, position) and *not* form.

Table 45

	S _H	N	O	D	A		
239				la _j	Ø _i	dije la verdad _i	I told her the truth
240				te _j	la _i	dije e _i	I said it to her
241				le _j	Ø _i	dije la verdad _i	I told her the truth
242	A María	se	le _k	Ø _j	Ø _i	enviaron los regalos _i a casa _j	They sent the presents home for her _k
243	El regalo		le _k			fue enviado	The present was sent for her _k
244	La carta		le _k			llegó tarde	The letter arrived late on her _k

3.4.3 Lower Benefactives

RAE (1973) considers that indirect-objects may be marked by *a* or *para*. This is motivated by similarity in meaning, whereby it is claimed that (245)=(246).⁴¹ As Maldonado (2000a) *i.a.* show, however, *para* profiles distal and projective meanings: to future time (247), to events yet to develop (248), or to event external participants, possibly not arriving (249). Whilst, DAT operates as *container* of (and is affected by change in) ACC, benefactives are merely *reference points*: the preposition *a* profiles affectedness, *para* merely indicates subject *intention* of contact/coincidence.⁴²

Table 46

245	Han traído un paquete para el director	The have bought a package for/to the director
246	Le han traído un paquete al director	
247	Lo quiero para mañana	I want it for tomorrow
248	Te lo repito para que entiendas	I'll say it again for you to understand
249	Se _i lo dieron [a José _i] para toda la familia, no para él _i	They gave it to J. for all the family, not for him

Datives using *a*, establish physical/mental contact with their object (250-251), whilst

41 Pottier (1971) treats datives and benefactives as the same functional category. Others see them as contrasts eliminated as benefactives “advance” (Perlmutter 1983) or are “incorporated” (Pool 1990) to dative markers.

42 Delbecque (1995) and Lewis (1989) for similar characterizations of *a* and *para*.

benefactive *para* denotes distance, leading to unacceptable results if used where such contact is inherent (250) or intended (251). Since benefactives indicate intention, they cannot determine the logical consequence of acts. In (252), *a* and *para* may alternate. A second clause may be logically consequential upon the first clause's transfer (*le...a*) (253), but cannot receive this reading in benefactive (254). Dative constructions establish links between participants, benefactives simply designate subject intentions, regardless of achievement. (255) deals with multiple *potential* recipients; since there is no knowledge of affectedness by those recipients, clitics are questionable. As distance increases affectedness diminishes. *Le* is simply inadmissible in (256) since the NP cannot possibly be considered affected.

Table 47

250	Le cepilló el pelo {a/para} Valeria	He brushed Valeria's hair
251	Le puso la falda (*para María, with contact reading)	He put her the skirt
252	Leí un libro a/para los niños	I read a book to/for the children
253	Les leí un libro a los niños...y se quedaron dormidos	I read a book to the children...and they fell asleep
254	??Leí un libro para los niños...	I read a book for the children...
255	Él (??les) escribía novelas para las damas de su época	He wrote novels for the ladies of his times
256	Él (*le) barre banquetas para el gobierno de la ciudad	He sweeps the streets for the city council

Semantic differences are reflected in syntax. Alarcos Llorach (1970) and Vázquez (1995) *i.a.*, note that fronting indirect-object PPs must be accompanied by dative clitics (257), i.e. valent datives must be filled, whether overtly or by $[-\text{SPEC}, \text{DAT}] = \emptyset$. Clitics with fronted benefactives, however, are ungrammatical (258). The *para*-phrase's referent is not a verbal argument, but rather stands outside the event. (259-261) illustrates how different construals of the same situation are directly coded into syntax: (259) the event is independent, but evaluated from the perspective, of the participant; (260) the subject performs the event with the external participant in mind (i.e. intention); (261) the event includes the participant who actually takes possession and is thereby affected.

Table 48

257	Al director _i , (le _i) han traído il paquete	[They brought the package to/for the director]
258	Para el director _i (*le _i) han traído il paquete	[They brought the package] for the director
259	Ella Ø hace un pastel para él	She bakes a cake for him
260	Ella le hace un pastel	
261	Ella le hace un pastel a él	

If X bakes bread for Y, Y may be present within the action and thereby possessor of the bread (DAT), or absent where X acts for Y's benefit (OBL) i.e. X carries out the event with the intention of giving the bread to Y at some future time. If the subject bakes bread for him/herself (X=Y), (s)he must logically be present within the action. The fact that the actual benefit is seen as a future event (i.e. possessing the finished product) is irrelevant. The subject is at all times the possessor of the bread whether as flour, dough or a loaf. It follows that reflexive benefactees are always DAT, whilst non-reflexive benefactees may be DAT or OBL according to context. There is no situation where OBL can be reflexive since this would involve being the subject of an action at which (s)he is not present (see §3.4.1).

Borer & Grodzinsky (1986) offer a syntactic diagnostic: possessor datives can be questioned, OBL cannot. (262-263) represent creation/destruction transitives employed as ditransitives, with datives construed as recipients, as in (260-261). Variation in acceptability of such datives depends entirely on the compatibility of verbal meaning and its object in ditransitive contexts (Leclère 1976:74). These are, therefore, also internal arguments.

Table 49

262	Paul a ouvert cette porte à Marie	Paul opened this door on/for Mary	French
	À qui est-ce que Paul a ouvert cette porte?	For whom did Paul open this door?	
263	Paul a fabriqué une table à Marie	Paul made a table for Mary	
	À qui est-ce que Paul a fabriqué cette table?	For whom did Paul fashion this table?	

Target datives must be active participants. Spanish clitics cannot substitute purely locative expressions. The benefactive reading (*for*, not *to*) of *a la señora* (264) highlights that she is not only a location but also positively affected by the chair's movement into her domain, not simply to her location. Datives can be coreferential, as subjects transfer objects into their own domain whether they are also locative targets (266), or not (265). Non-affected locatives are not datives. *Le* cannot be linked to *mesa* (268), nor made reflexive. (266) is the reflexive counterpart of (264), not of (267).

Table 50		
264	Le _i acercó la silla a la señora _i	He pulled the chair up for the lady
265	Se compró una falda	She bought herself a skirt
266	Se acercó la silla	She pulled the chair up for herself
267	Acercó la silla a la mesa	He pulled the chair up to the table
268	*Le acercó la silla a la mesa	

Delbecque & Lamiroy (1996) treat verbs like *unir* as part of the *añadir/aplicar/asociar* type which take dative complements, considering that such verbs “can also be construed with the preposition *con* provided the correspondence is conceived as coincidence.” For *añadir*, affectedness occurs in an “incorporative” sense (269), however, with *con*-verbs (270) neither entity undergoes changes-of-state; they merely become coincident in concrete/abstract space. Even with *a* (which is marginal), *le(s)* is precluded (271). The relationship between entities remains symmetrical and, therefore, unaffected. Equally, verbs profiling subject movement to locative goals (e.g. *acceder*, *acudir*) cannot take dative clitics. The subject's arrival denotes coincidence not incorporation and hence does not affect (272).

Table 51			
269	Se le añade azafrán al arroz	One must add saffron to the rice	(Maldonado 2002a)
270	Alió indios con meztizos	He united Indians with Mestizos	
271	(*Les) alió indios a Meztizos	He united Indians to Mestizos	
272	(*Le) accedieron al senador para...	They went to see the senator in order to...	

Individuals can be beneficiaries within, and by virtue of, an event. The two categories must be kept separate, otherwise the syntactic properties discussed above become merely stipulations. In a structure which has separate places for each, iconically representing those relationships, such phenomena emerge naturally. They do not even need to be mentioned.

3.4.4 Emphatics

Emphatics highlight structural distinctions between upper- and lower-fields. When object arguments are emphasized, emphatics must agree with that object (273). Whilst addition of *mismo* is obligatory with reflexives (274), including benefactives (275), non-arguments i.e. possessive DAT with monotransitives (276) or OBL (277) cannot take *mismo*, even if reflexive. (278) may only be read as benefactive (279), equivalent to (275). Thus, (280) cannot be interpreted as possessive; only benefactive (281). Without *mismo*, it may be read as default possessive (282) or benefactive (283), with *a sí mismo* forcing benefactive readings (281).

Table 52

	N	O	D	A		Intended Reading
273				lo	le lavé a él (mismo)/ella (misma)	I washed him Direct Object
274			se	Ø	lavó a sí mismo	He washed himself Indirect Object
275			me	lo	lavé a mí mismo	I washed it for myself Benefactive
276			*le	Ø _i	comí la paella _i a él (*mismo)	*I ate his paella *Possessive
277		*le				*I ate the pie on him *Malefactive
278			*se	Ø _i	lavó el coche _i a sí mismo	*He washed his own car *Possessive
279			se			He washed the car for himself Benefactive
280			*se	Ø _i	comió la paella _i a sí mismo	*He ate his pie *Possessive
281			se			He ate the pie for himself Benefactive
282			se	Ø _i	comió la paella _i a sí mismo	He ate his own pie Possessive
283						He ate the pie for himself Benefactive
284	se			Ø _i	comió la paella _i a sí mismo	He ate up the paella Agentive
285					comió mi paella _i	He ate up my paella
286				lo	hizo él mismo	He did it himself Subject Emphatic
287				Ø	él mismo limpió el coche	He himself washed the car

SE may also produce agentive readings (284). Unsurprisingly, subjects are not emphasized with object *a sí (mismo)*, but with nominative forms (286-287), so (280) cannot be read as emphasizing SE_{NOM} and cannot clash with the forced benefactive reading. Where verbs, e.g. those of consumption, tend to take SE_{NOM}, possessive readings are forced through possessive adjectives (285). Thus, despite minimal signals, default interpretations will usually lead to correct interpretation, whilst any vagueness has specific resolutions, if and when greater precision is required.

3.4.5 Putative PCC-Breaches

There is some confusion concerning object-reflexive usage. This section briefly shifts to Italian, since proscription of *OBL+SE makes these far less common in Spanish.

Some see (291, Cardinaletti 2008:78) as a PCC-breach. This, however, is a misreading of such sentences. *Rivolger+si* represents two constructions. In (288-291), *si* is reciprocal indicating shared ownership/destination of the explicit (289) or implicit (290) object *parola*. In (291), *in inglese* shows that the construction is transitive i.e. words (inherent ACC) of the subject (*si*) are being directed to some place (*mi/gli*). Neither *a te* (290) nor *mi/gli* (291) are DAT, since that role is taken by the possessor. In (292-294), *rivolgersi* is a verb of disposition (\approx *girarsi*) taking *a*-phrases indicating direction i.e. place (293) or person (294, *a lei*). The distinction between the two phases (turning and subsequent actions) is highlighted in the translation ‘go and’. There is no transfer except in the *dicendo* sub-clause (294). In (292-294), the subject turns himself (SE_{ACC})⁴³ or ‘becomes’ (SE_{MID}) oriented towards someone/something. The other participant is not a verbal argument in (288-294), but a situational affectee/place (OBL), as reflected in its restricted use in complex clauses (295-297, Cardinaletti 2008). In clitic-

43 This would most likely take a passive reading ‘was turned to’ which is not intended.

climbing configurations,⁴⁴ lower clitic-fields attach to infinitives whence they may climb to the modal verb's lower field. If this were *mi*_{DAT}+*si*_{ACC}, both could cliticize to the infinitive, however, *mi* cannot; it may be applied to the whole verb complex as a complement (*a me*_{OBL}) or appear as the verbal complex's OBL. Conversely, *si* as a verbal argument of the lower event, remains attached to the infinitive, or raises to the matrix verb's DAT position.

Table 53

288	Rivolgere la parola	To address somebody (transitive)	Italian
289	Non si Ø _i rivolgono più la parola _i	They are no longer on speaking terms They no longer address their speech to each other	
290	Non mi Ø _i rivolgevo Ø _i a te	I wasn't speaking to you I was not directing my words to you	
291	Mi/gli si è rivolto in inglese	He addressed {his words/himself} to me/him in English	
292	Rivolgersi a (per informazioni)	Go and see/go and speak to	
293	Rivolgersi all'ufficio competente	To apply to the office concerned	
294	Si rivolto a lei dicendo...	He turned to her, saying...	
295	Mi si è rivolto in inglese	He addressed himself (i.e. his words) to me in English	
296	*Vorrebbe rivolgermi in inglese	He would address...	
297	Vorrebbe rivolgersi a me in inglese	He would [address himself in English] to/on/for me	
298	Se la avvicina	He draws it to himself	
299	Il treno si avvicinava alla stazione	The train drew near to the station	
300	Si avvicina l'inverno	Winter draws near	
301	Mi si avvicinò un mendicante	A beggar came up to me	

Many verbs follow identical patterns. In contrast to transitive (298) with its object and reflexive recipient, (299-300) are non-active. These are not passives: subjects are not effected by an external force. Rather, *SE*_{MID} indicates a developing internal COS of approaching a place; explicit (299) or implicit (300, *here-and-now*). In (301), *mi* is not a recipient/possessor, but an orientation (referenced via a participant) and/or an experiencer/affectee of the event. Semantically and syntactically, *mi* is OBL. Such cases are not PCC breaches.

44 Note in non clitic-climbing environment, infinitives may also carry upper clitics.

In all these cases, a simple model which clearly separates OBL~DAT and place~possession ([±E]) properties is able to express the range of meanings and functions found in real-life usage. Speakers are able to express their ideas directly through an iconic model of their construal and reasonably expect the listener to be able to parse and understand that message, without learning complex rules. What are presented as problems or stipulations in other approaches, simply emerge from the proposed model.

3.4.6 Conclusions

Participants referenced by OBL are coincident with the event (within the speaker's construal). Position tells us that they are outside the event, context tells us whether they act as experiencers ([−E]) or evaluators ([+E]). Excepting Romanian (§7.4.4), lack of DAT_[+E] clitics, produces DAT~Ø alternations, exploited in dative-doubling languages to compensate for lack of locative clitics.

Table 54

$f_x(D\ A)$	$\rightarrow [E\ D\ A]$	Coincidence of D and A	within event	[+E]	Ø
	$\rightarrow [E\ D(A)]$	Possession by D of A	within event	[−E]	DAT

$f_x(O\ [E\dots])$	$\rightarrow O\ [E\ A]$	Coincidence of O and event effecting A	external to event	[+E]	OBL
	$\rightarrow O\ ([E\ A])$	Possession by O of event effecting A	external to event	[−E]	OBL

	N	O	D	A
Affectedness ⁴⁵	±	✓	✓	✗
Participation	✓	✗	✓	✓
Truth Conditions	✗	✗	✓	✓

OBL's appearance with intransitives where verbs only select subjects, and inability to be emphasized or coreferent indicates that it is not a verbal argument. Whilst datives differ from

⁴⁵ SE_{NOM} may indicate 'satisfaction' (§4.7.2).

NOM/ACC by virtue of affectedness, OBL is distinct by virtue of non-participation. Whilst DAT/ACC are directly involved in events, where modification will change clausal truth conditions, upper-field clitics introduce non-truth changing elements, contributing to meaning at the pragmatic level, highlighting subject and third-party perceptions of the event.

These two datives are semantically, syntactically, and *positionally* distinct. Without this understanding, pairs of dative clitics cannot be accommodated by person-models, except in the rare and fortuitous case that they happen to meet templatic requirements. With two distinct positions, it is possible to explain when they do and do not appear in each function and cover the full range of data with recourse only to non-clitic specific RND and the PCC for exclusions.

3.5 Communication Theory and Clitic Patterns

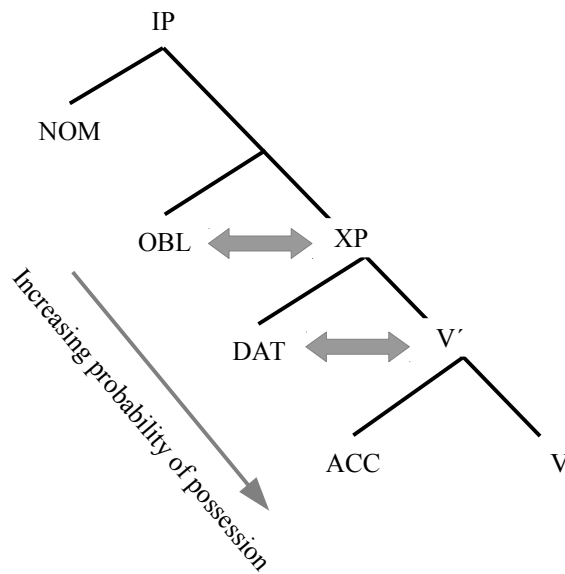
This section places the proposed structure in relationship to Cognitive Linguistics' three main tenets of *iconicity*, *compositionality*, and *interpretation*, whilst highlighting the positive value of *vagueness* in natural language.

3.5.1 Signalling Relationships

It is clear that OBL/DAT cannot be merged merely because the clitics in these positions take the same forms. Besides being able to appear together, they have different semantic/syntactic functions; both are *affected*, but what *affects* and *is affected* are different.

$$[_{IP} \text{ OBL} \leftarrow [_{XP} \text{ DAT} \leftarrow \text{ACC}]]$$

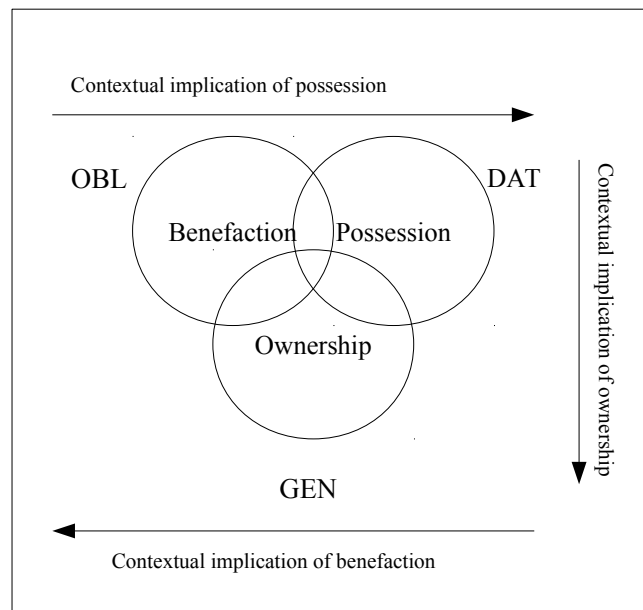
A clitic's function is *not* directly to express *any* particular set of properties, but rather signal *significant relationships* between participants within an event (DAT), or non-participant and event (OBL). The nature of these relationships is *not* clearly defined, merely their existence.



Romance clitics do not encode positive, negative, vs. static relationships. There is no surface distinction between allative/ablative/genitive relationships between event participants; all take dative-form under DAT. There is no distinction between benefactive/malefactive/experiencer relationships between non-participants and the event; all take dative-form under OBL. ‘Direction’ is determined by the verb and situation. Similarly, possession is *inferred* from the presence/absence of these clitic signals, the particular context (i.e. knowledge of the participants), and shared world knowledge (i.e. what is more likely).

These semantic relationships are matched by the syntactic model, with likelihood of possession increasing as the signal approaches the possessum. OBL marks relationships to the event and less directly to ACC; affectedness *may be* due to possession. DAT marks

relationships directly related to ACC; affectedness *probably is* due to possession. In either case, possession *may* imply ownership. Conversely, possessive adjectives within ACC's DP *must* indicate ownership, but only *probably* possession. This can be seen in the tendency to read isolated OBL as DAT; there is nothing to 'distance' it from ACC.



Like 'direction', possession is not expressed in surface-forms. The three closely related concepts of intended/actual possession (DAT), external benefaction/malefaction (OBL), and ownership (possessive adjective) form an overlapping domain, in which more than one property may be true of the object (ACC). The property indicated by each syntactic unit is distinct i.e. *significance* at the level of participant, non-participant, or ownership (i.e. outside of the contrual). The listener infers related properties from expectation and/or context.

3.5.2 Parsing and Efficiency of Communication

Two adjacent syntactic positions with identical surface forms to express their referents is problematic when considering surface sequences *in vacuo*. Whilst it is easy to extricate these functions when they appear side-by-side, one is reliant upon expectation and context to

interpret isolated datives. It is important to remember that clitics represent old information (i.e. interlocutors have expectations) and that OBL forms are most common in spoken language (i.e. between interlocutors who have built shared context). The analyst's difficulty arises from snatching surface-forms out of context; in real life, this problem does not arise. Even cases which are technically vague are rarely misleading as default interpretations come into play. In cases where lack of communication might ensue or something outside the norm is intended, important details can be emphasized/denied through additional adjuncts. The speaker knows when these additions are necessary for the listener because of their shared view and selects the elements necessary to compose his intended message appropriately.

Indeed, interlocutors do not expect expression of all properties. The *vagueness* (often confused with ambiguity) which plagues analysts is, in fact, a sign of linguistic *efficiency*. It has been a central tenet of Communication Theory since Saussure (1916) that language cannot transfer all data. Each speaker construes a situation and presents *sufficient* data for the listener to re-build it in his mind from the *minimal* pointers provided in speech and shared knowledge of context. The speaker need only signal relationships as significant to the communication by inserting the appropriate clitic. The shared inference engine will (99% of the time) provide the full picture.

Contra most analysts' implicit view, there is an expectation (indeed requirement) for natural language to display vagueness. Strozer (1976:156) reports the following real-life exchange:

A: ¿Le lavaste el coche a tu papa?
B: No, me lo lavé *a mí mismo*.

Intended: Did you wash your father's car?
No, I washed it for myself.

B has misinterpreted *A*'s question as “did you wash the car *for your father?*” The interlocutors view the same situation from different perspectives, and therefore interpret identical signals differently, based upon their initial biases; *A* is concerned with the father, *B* with himself. In a longer conversation, a shared viewpoint will develop and vagueness will reduce. However, if confusion arises, speakers simply add the necessary extra material to make things clear; but only when necessary.

By virtue of such automatic inferences, increased explicitness signals variation from the norm. In ‘default’ contexts such explicitness becomes misleading to the listener. The gap between the correct default interpretation (denied by over-specification) and an alternative (demanded by inappropriate levels of specificity) causes a psychological dissonance often referred to as *ungrammaticality*. As illustrated, most *unacceptable* usages are reasonable given an appropriate context, and therefore, should not be the subject of ‘rules’ to ban them. In these cases, *ungrammatical* simply means *inappropriate to context*. Their inappropriateness is precisely because the listener *expects to interpret* the spoken message from context and *minimal* signals.

Returning to the Cognitive Linguistic approach discussed in §1.4, semantic properties are *iconically* reflected in structure, which guides *interpretation* through its inner→outer sequence. Structure tells the listener that more than one option is available, whilst default strategies (over-ridden by explicit data) lead to selection of an appropriate schema. In some cases, there will be more than one possibility and limited (and correctable if necessary) differences in understanding will ensue. It is in such limited, but still effective, miss-communication that historical change finds its means.

3.6 Conclusions

This chapter has outlined the existence and communicative need for the distinction between OBL~DAT and the need for [\pm E] in both. The multiplicity of uses examined underlines the need to distinguish form from function, and the important role of inference which can only take place in terms of the sequential structure in which these clitics are presented. In the following chapters, we extend these ideas to reflexives (Chapter 4) and non-personal clitics (Chapter 5).

4 THE UBIQUITY OF SE

This chapter explores reflexive clitic *forms*. In most Romance languages, the same forms are used for all cases, and for reflexive and non-active uses, highlighting the need to separate form from function. We argue that those functions can be identified through differences in syntactic/semantic usage and are more numerous than previous analyses have allowed for. We express this range in terms of ‘case’ (i.e. position in the clitic field) and [$\pm E$] (representing the disjoint vs. subset distinction), allowing us to clearly identify the full range of impersonal, reflexives, and non-active concepts of middle, passive and anticausative. We hope to show that the wider range of functions which the model predicts do indeed exist and, moreover, are necessary for languages to be able to express the full range of meanings for which these forms are employed.

4.1 Introduction

Reflexive pronouns, particularly in 3-person, have proved problematic for all approaches to clitics. As well as replacing coreferent (in)direct-objects, they may also indicate non-active voice, impersonality, and volition. Their heterogeneous range of functions (“*polivalencia*”, Di Tullio 1997; “*carácter cameleónico*”, Otero 1999) has led to equally bewildering arrays of classifications. However, it is crucial to gain an understanding of this ‘system’, if we are to defend the approach outlined in Chapter 2.

Since some usages are restricted to 3-person, investigations tend to revolve around that form. Unlike other persons derived from Latin personal pronouns, the 3-person form derives from IPSE. Since it has developed several forms (e.g. Spanish *se*, Italian *si*), we follow the convention of referring to it as SE regardless of language. This also has the advantage of

reflecting key typological distinctions between Romance reflexives and those of languages such as English, often referred to as the SE~SELF distinction (cf. Reinhart & Reuland 1993).

4.1.1 The Problem

Spanish SE displays the greatest number of uses in any one language, including almost every use found in any Romance language. Contreras (1964) proposes 13 types of SE (1), illustrating not only its multi-faceted nature, but also the difficulty of achieving even descriptive adequacy. While most authors combine cases into larger sets, others argue that other significantly different uses are missing e.g. anticausative and intransitive impersonal.

The consensus divides cases between “true reflexives” and those “only of form” (Alonso & Henríquez Ureña 1971:104-105), where the latter uses do not imply any sense of ‘reflecting back’ onto the subject (“*cuasi-reflejas*”, Bello & Cuervo 1960:457). The latter heterogeneous group are variously sub-categorised: Montes Giraldo (2003) has 12 categories, Lázaro Carreter (1964) 9, and Hernández Alonso (1966) 6. Unfortunately, there is no agreement on terminology, and the same descriptive label may be used for different or overlapping concepts across authors.

Table 55

1	Reflexive-SE	Se lava	He washes himself
	Reciprocal-SE	Se observan	They watch each other
	Passive-SE	Se firmó el acuerdo de paz	The peace treaty was signed
	Impersonal-SE	Se aplaudió a los artistas	The artists were applauded
	Aspectual-SE	Se durmió	He fell asleep
	Diaphasic-SE	Se murió	He died
	Lexical-SE	Se fue de su casa	He went away from his home
	Affective-SE	Se bebió un vaso de vino	He drank up a glass of wine
	Morphological-SE	Se arrepintió	He repented
	Dialectal-SE	Se enfermó	He got sick
	Narrative-SE	Éra-se una vez un rey	Once upon a time there was a king

Many consider *dialectal-SE* and *diaphasic-SE* as evidence of non-pronominal *lexical-SE*. *Narrative-SE* (retained from Old Spanish for stylistic purposes, particularly in fairy-tales) is rarely mentioned in other studies, although it has major significance for impersonal/passive uses (§4.6). Finally, it should be noted that many works consulted make valued judgements concerning acceptability of particular constructions, discarding cases considered erroneous (see particularly §4.6.7). An adequate model, however, must reflect actual usage even if it offends grammarian sensibilities, particularly when deprecated forms are often norms in related languages, and even in earlier stages, or contemporary dialects, of Spanish itself.

4.1.2 Unity vs. Diversity

For Spanish, Monge (1955) traces processes by which all modern SE's functions derive by progressive extensions of possibilities already extant in Latin. Originally restricted to animates, Latin reflexives extended to inanimates by the first centuries AD, with expressive/emphatic function. For Monge, SE became merely a grammatical function employed as an 'intransitivizer', whilst retaining a sense of subject participation i.e. 'middle' value. Passive-SE is found from the earliest Spanish texts, constituting a further stage of grammaticalization. It was less frequent with animate subjects, possibly explained by potential confusion with reciprocal/reflexive readings e.g. *se mataban los cristianos* (Fernández Ramírez 1964:283; RAE 1973:§3.5.6b). Potential confusion joins with the Spanish tendency to syntactically distinguish animate objects with personal-*a* and sees the rise of constructions such as *se mataba a los cristianos* (§4.6). SE with intransitives constitute the last phase of evolution. §4.6.6 discusses a further stage: development of true 'impersonal' SE.

Structuralists, functionalists and generativists alike, consider SE as primarily a reflexive pronoun, which has developed an additional grammatical function as ‘intransitivizor’ (e.g. Babcock 1970; Álvarez Martínez 1989; Di Tullio 1997), sometimes termed ‘transpositor’ (Carratalá 1980:216-218; Martínez 1981; Alarcos Llorach 1994:§7). In non-reflexive cases, SE is seen as blocking the appearance of second actants; “diátesis recesiva” (Tesnière 1994:473). Fernández Ramírez (1986:399) discusses ‘neutralization’ of transitive verbs converting to them into “*verdaderos verbos intransitivos*”. No definition is given of the ‘intransitivizor’, however, merely descriptions of its activities.

Table 56

2	<La> acordaron <la paz>		≈Resolvieron de común acuerdo
3	Se acordó de memoria		≈Recordó
4	Entiende	los negocios	≈Comprende...
5		de negocios	≈Sabe de...
6	Reparó	los baches	≈Arregló...
7		en los baches	≈Notó, Miró con cuidado...

Several points bring the basic concept into question. There are cases where presence of SE does not eliminate the actant but rather causes its expression as prepositional-, not direct-, object. Alarcos Llorach (1970:217) sees SE’s function ‘purely’ to signal this semantic change and, by taking direct-object position, relegating true objects to supplementary phrases (2~3). However, similar alternations producing similar semantic changes are observable without SE (4-7). Thus, SE is not what is ‘blocking’ the object.

Whilst accepting the effect of ‘suspending’ verbal valency, Gutiérrez Ordóñez (2002) argues that, even in cases without such complications, SE cannot be an ‘intransitivizor’ because it

affects the subject as event cause and only on its suppression does the object raise to take subject position. Syntactically, the process is closer to passivization than intransitivization. This raises the question of how to ‘intransitivize’ already intransitive verbs. Some opine that these represent causative~inchoative alternations, hence the suppressed argument is the cause; however, cause is not always relevant, nor always suppressed (§4.3.2). Finally, several authors have noted that impersonal-SE simply does not fit such simple dichotomies (§4.6).

Despite being treated as having ‘no syntactic function’ (“*un mero componente verbal*”, Gómez Torrego (1992:18), this ‘intransitivizer’ is found with heterogeneous sets of verb classes producing an amorphous collection of semantic effects, which cannot (in our opinion) be attributed to a single ‘transpositor’. In addition to marking specific (although often subtle and difficult to evaluate) semantic changes, use of SE in a given context is (we shall argue) circumscribed by, and interacts with, the syntax of the whole predicate, indicating that each SE has particular syntactic (as well as semantic) properties. Indeed, most analysts further subdivide uses of SE by various means, only achievable in terms of syntax. We, therefore, reject the concept of ‘intransitivizer’, working from the premise that each usage can be identified through its syntactic function and semantic effect.

4.2 Reflexive SE

This section provides a formal basis for the ‘true’~‘only-of-form’ division, with some unexpected consequences.

4.2.1 Reflexive Functions

Lidz (1997, 2001) and Reuland (2001, 2005) distinguish two reflexive types.⁴⁶ In pure-

⁴⁶ Hebrew (Doron 2003) and Kannada (Lidz 2001) have distinct pure- and near-reflexive surface-forms.

reflexives (8), the two arguments are identical in the world and semantic representation. In near-reflexives (9), the second argument is a function upon the first, returning an entity related to that argument. The near-reflexive function $f(x)$ allows the antecedent and anaphor to be the same world entity but does not require it, as found in Tussaud contexts (Jackendoff 1992). In ‘X sees himself’, the reflexive references an antecedent distinct from the subject (i.e. statue vs. person). In French, near-reflexive interpretation is possible in Tussaud contexts with reflexives (10)⁴⁷ and reciprocals (11). Reuland (2005) associates $f(x)$ with conditions of near identity, where the object ‘stands proxy’ for its subject. Thus, a statue of X may stand proxy for X, but not a book about X.

Non-Tussaud contexts, however, require different definitions of $f(x)$. Ruwet (1972b:88) notes that (13) is not the reflexive of (14), which is non-existent, but rather means (15). The object is understood as P.’s ideas/opinions; a relationship akin to metonymy, not near identity (Labelle 2008 for a similar analysis). Importantly, these verbs cannot be interpreted reciprocally; (16) requires that each subject explains their *own* behaviour, in parallel. *Se* is subject-coreferent possessor, and hence licenser, of an ellipsed accusative. Applying this to Tussaud contexts, (10) becomes not ‘admires himself’ where $f(\text{himself}) \rightarrow$ ‘statue’ (close copy), but ‘admires *his...*’ where the object is drawn from *his* possessional domain within each context (e.g. Tussaud or not).⁴⁸ Whilst *l’un l’autre* (11) modifies the manner of the event forcing reciprocal readings but leaves argument interpretation to $f(x)$, the addition of *lui-même* to (10) over-rides $f(x)$, forcing a direct ‘self’ (12) interpretation (see object contrast, §4.2.2).

47 Rooryck & Vanden Wyngaerd (1999) prefer *se...lui-même* here; Labelle (2008), the opposite.

48 See also Cognitive Linguistic’s active zone (Langacker 1987:271–4, 1993:29–35). The intended referent of *We all heard the trumpet* does not match the semantic content of the object argument, but is in experiential contiguity to it (Traugott & König 1991:210–2) i.e. part of its abstract domain.

Table 57

8	$\lambda x[P(x, x)]$	Pure-reflexive
9	$\lambda x[P(x, f(x))]$	Near-reflexive (1)
	Au Musée Tussaud,...	At the Tussaud Museum,...
10	J a pu s'admirer	...J. was able to admire himself (=his statue)
11	P et M ont pu s'admirer l'un l'autre	...P. and M. could admire each other (=each other's statue)
12	J a pu s'admirer lui-même (dans la glace)	...J. was able to admire himself (=his image) in the mirror
13	P s'est exprimé avec clarté	P. expressed himself clearly
14	*P a exprimé Paul avec clarté	...P. clearly
15	P a exprimé ses idées avec clarté	...his ideas clearly
16	J et M se sont expliqués	J. and M. explained {their own/*each other's} behaviour
17	J et M se téléphonent	J. and M. telephone/made a call to each other
18	J Ø se _{ACC} lave	John washes himself
19	J se _{DAT} Ø lave	...his (self)
20	Las manos, J se las lave	The hands, John washes (his) them
21	Ils se peignent (les cheveux)	They comb their hair
22	$\lambda x[P(x, x)]$	Pure-reflexive
23	$\lambda x\lambda y[P(x, f(y)) \text{ where } x=f(y)]$	Near-reflexive (2)
24	$\lambda x[R(x, x)]$	'Closed', 1 semantic argument
25	$\lambda x\lambda y[R(x, y) \wedge x=y]$	'Open', 2 semantic arguments

This approach uses the same mechanism for both contexts, and reflexive possession in general. Furthermore, it explains restrictions on 'reflexive verbs'; $f(x)$ is a part/whole relationship, most easily inferred in cases such as personal grooming, where effected objects are simultaneously part of the subject (19). Such reflexives are not unaccusatives,⁴⁹ since dative reflexives/reciprocals exist (17), including ones with accusatives (20). For some verbs, objects are inherent (21, hair), others default to 'self' (18-19), but parts may be individuated (20).⁵⁰ Where possession is shared reciprocity is inferred with uniplex (17) vs. multiplex (16) interpretation generated by $f(x)$ from context and inherent verbal semantics. In all these cases, *se* fills the appropriate argument slot, hence being obligatory whether reflexive/reciprocal pronouns are present or not (see case contrast, §4.2.3).

49 See §4.2.5 for arguments against this simplistic equivalence.

50 Whilst direct- (18) and meronymic- (19) reflexives were common in Old French (Kemmer 1993:153-62), most metonymic reflexives (13-16) arose only in Modern French (cf. verbs listed in Hatcher 1942:155-6), pointing to an expansion of the boundaries of what is considered possessable.

This more detailed definition of reflexivity (23) matches distinctions between ‘closed’~‘open’ predicates (Sells *et al.* 1987, *i.a.*). ‘Closed’ predicates (24) possess unique variables saturating two thematic roles, whereas ‘open’ predicates imply two semantic arguments where one refers to the same entity as the other (25) but where object interpretation is not necessarily bound to that of the subject. In Sells *et al.*’s terms, ‘closed’ predicates are semantically intransitive (one variable), ‘open’ predicates are semantically transitive (distinct variables).

4.2.2 Contrastive Pronominals

Based upon participant contrast, Labelle (2008) argues that French reflexives are ‘open’ predicates (similarly Bruening 2006 for reciprocals). In (26/27), *lui-même* places focus on the object without intonational prominence, by overtly contrasting it with other potential objects. The background is obtained by replacing focused objects with a variable ranging over potential entities (Rooth 1992; van Heusinger 2004; *i.a.*). Thus (27) asserts (28) against background (29) in which event goals might be different from *ministre*. Contrastive focus acknowledges the possibility that the object might be distinct and, therefore, predicate interpretation requires positing distinct agent~goal variables (Rooryck & Vanden Wyngaerd 1999). French reflexive/reciprocals are, therefore, ‘open’ predicates (and semantically transitive), since ‘closed’ predicates exclude this possibility.

Table 58

26	Le ministre se copie lui-même	Direct reflexive
27	Le ministre se parle à lui-même	Indirect reflexive
28	$\lambda e[\text{say-to}(e, \text{ministre}) \wedge \text{Agent}(e, \text{ministre})]$	Assertion
29	$\lambda x \lambda e[\text{say-to}(e, x) \wedge \text{Agent}(e, \text{ministre})]$	Background
30	J ne se rase pas	J. does not shave
31	J ne se rase pas lui-même	J. does not shave himself
32	J ne s’est pas dénoncé lui-même	J. did not denounce himself
33	J [a acheté la chemise] lui-même	J. has bought the shirt himself
34	Jean [la connaît] elle	Jean her knows her (Kayne 2000)
35	Les enfants se sont suivis	The children followed each other
36	Pierre et Jean se sont écrit l’un à l’autre	Pierre and Jean wrote to each other

In (27), *lui-même* is dative-marked, but (26) is ambiguous, defaulting to interpretations as accusative and introducing object contrast. However, *lui-même* may also be subject-oriented, opposing actor to other potential actors. Rooryck & Vanden Wyngaerd (1999) note that in (30), no shaving occurs (J is bearded), while in (31), J is shaved but is not the actor. An object contrast reading is also possible in (31), where J shaves someone else. While the preferred interpretation of (26) is object-oriented and of (31) as actor-oriented, (32) is compatible with interpretations in which J was denounced (e.g. in prison) by others, and where J denounced his friends, not himself (he is free). Default readings are derived from world knowledge; people (other than barbers) tend to shave themselves, whilst denunciations work both ways. In contexts where the object is known (33), *lui-même* can only serve as an ‘actor-oriented intensifier’. Crucially, nominative case is equally as contrast-able as dative and accusative (see case contrast, §4.2.3).

Table 59

37		38		
a.	The boys slapped themselves	(each only his self)	Les étudiants se sont frappés	aux mêmes
b.		(mixed)		Ø
c.	The boys slapped each other	(each only another)		l'un l'autre

Reflexive situations fall into three categories: fully reciprocal, fully reflexive, or mixed. 61.4% of languages (Heine & Miyashita 2008) follow the English pattern (37). Each pronominalized sentence is specific to its context; neither is available for mixed situations. Romance follows (38) with *se* in all contexts. This breadth of readings is not (as sometimes described) ambiguity/polysemy (Heine & Miyashita 2008; Gast & Haas 2008; Maslova 2008), but vagueness (Cable 2014). *Se*-reflexives are not ambiguous between (38a~38b), but possess a single, weak interpretation encompassing all situations e.g. in (35), any combination of ‘following’ is allowed. Details (often irrelevant) are inferred from context or highlighted

when required using appropriate adjuncts.

Given the possibility of *se*-reflexives with/out adjuncts, it must be *se* that fills the required argument position, indicating broad ‘reflexive’ readings. Adding pronominals merely *enriches* context, highlighting specific portions, *without* interacting with verbal valence. This is confirmed by subject-contrast (33), where the pronominal cannot be in argument position, and since there is no ‘reflection’ (the verb has its own object), no *se* appears. Reciprocal pronominals follow similarly. Following Déchaine & Wiltschko (2004) *i.a.*, the structure of *l’un l’autre* is [_{distributor} l’un [*e* [_{reciprocator} l’autre]]], where variable *e* is bound by *se* and co-indexed with the plural subject and *l’un+l’autre* refer to members of the set denoted by the subject. In subject/object reflexive/reciprocal contrast, pronominals stand outside of VP arguments, like *elle* in (34).

4.2.3 Case

(39) expands upon (37-38), highlighting the importance of case, which as (17-20) illustrate, must be taken into account even when no other argument is present i.e. SE_{DAT}~SE_{ACC} are syntactically distinct despite their syncretic forms, matching distinctions shown by non-reflexive clitics which maintain separate forms in most languages.⁵¹

Table 60

39	N...	D	A	...	[SUBJECT	[OBJECT]]	...	Contrast	Case	Interpretation	e.g.
	Ø		SE		[lui-même]		Object	ACC	Reflexive	(26)
	Ø	SE			[à lui-même]			DAT		(27)
	Ø		SE		[None	ACC	Mixed	(35)
	Ø	SE			[DAT		(17)
	Ø		SE		[l’un	[l’autre]]		Mutual	ACC	Reciprocal	(11)
	Ø	SE			[l’un	[à l’autre]]			DAT		(36)
	Ø				[lui-même	[]]		Subject	NOM	Emphatic	(33)
	SE				[None		‘Expressive’	§4.7.2

⁵¹ This is historical accident. Romanian maintains $\text{și}_{\text{DAT}} \sim \text{se}_{\text{ACC}}$, whilst Gascon (§6.5.4) has no A/D distinction.

As noted in (§4.2.2), nominative is equally as contrast-able as dative or accusative. As with objects, *lui-même*'s function is contrastive and not reflexive. Since $f(x)$ can return x (§4.2.1), all reflexives are 'near' expressing different 'views' of the subject. There is no reason to presume that NOM cannot also take 'reflexive' functions, i.e. appear as SE. Transitivity denotes energy leaving the actor, entering the outside world, and (in)directly affecting participants in a given role. When that role is effector, it is logically possible to talk about 'nominative reflexives'.

Table 61

40	[FR]	Je me la bouffe	I gobble it up (Babcock 1970:65)
41	[SP]	Juan se lo comió todo	John ate it all up
42	[IT]	Gianni se lo mangiò tutto	
43	[SP]	Se te me lo llevó	He took it away from me on you (against your wishes)

In fact, such forms are found in most Romance languages with varying degrees of acceptability (40-42), generally introducing an element of subject 'intent' and/or 'satisfaction' with event completion. SE_{NOM} can be confused with other uses, but (43) shows that it can only be nominative since all other positions are simultaneously filled.⁵² An understanding of this category of clitics is developed in §4.7.2, once we have laid out the full range of uses of SE with which they contrast.

4.2.4 Emphatics

Spanish shows similar usage, although with different case-marking, due to Spanish employing personal-*a* with ACC_[+ANIM].

Table 62

	NOM	ACC	DAT
French	lui-même		à lui-même
Spanish	él/sí mismo	a sí mismo	

⁵² §7.5.1 for discussion of these rare and complex four-clitic clusters.

Otero (1999:1431-62) argues that *sí (mismo)* is the only Spanish reflexive pronoun because, unlike SE, it uniquely constitutes an anaphoric subject-reference permitting antecedents. Fernández Ramírez (1986:76-77), however, provides (44-45) where *sí (mismo)* refers to non-subject elements, and conversely (46-48) where, coreferent *sí* is not commonly understood as reflexive. Moreover, in colloquial usage, *sí* is often interchangeable with *él/ella* which, although referencing the subject, hardly qualify as reflexive (46-48). In short, *sí (mismo)* is not an effective test of reflexivity.

The central property endowed to predicates by *sí (mismo)* is [+intent].⁵³ Object-reflexive forms are neutral (49), allowing reflexive (*X hizo algo que afectó a X*) or non-active readings, often seen as accidental (*algo le sucedió a X*). Readings may be forced by emphatics (50) or adverbials (51-52), but cannot be mixed (53), or duplicated (54). Inherently agentive and non-accidental verbs e.g. *suicidarse* invalidate use of [–intent] adverbs, whilst applying additional [+intent] material leads to awkwardness (55), as with non-reflexive verbs (56). Conversely, although “*un tanto forzada, masoquista*” (Di Tullio 1997:174), verbs of physical damage can be read as subject controlled (57). The usual reading is [–control] with subject as patient and agentivity is irrelevant e.g. ‘he got his leg broken’. Adding *sí (mismo)* introduces intention (like the *para*-clause) denying this possibility and leaving only masochistic readings.

A sí (mismo) matches other circumstantial complements. The *a* is not personal-*a* introducing animate objects, but a simple preposition introducing an adverbial manner phrase, like *a mano*, aligning it with the full range of such phrases introduced by other prepositions. This

53 Van Valin & La Polla (1997:392-417) show that “coreferential reflexive constructions” as found in English (which seem to be the source of Otero’s conception of reflexivity), possess very different properties from Romance “reflexive clitic constructions”, particularly in terms of their representation of agentivity.

approach solves a major problem. If *a sí (mismo)* and its equivalents were reflexive pronouns, it would represent clitic-doubling, which is acceptable in Spanish but banned in most languages. Under the current view, its equivalent (*lui-même* etc) would be equally circumstantial and thereby grammatically acceptable along side clitics in all languages. Pederson (2005) provides evidence of historical change of meaning from reflexive to emphatic for Spanish, and Zribi-Hertz (1982) for French.

Table 63

44	En todas las casas están los solares de sí mismas	Spanish
45	Divertir es apartar a cada uno de sí mismo	
46	Pedro logró los objetivos por {sí/él} mismo	
47	María confía en {sí/ella} misma	
48	Juan tiene muchas personas detrás de {sí/él}	
49	P. se quemó _[0intent]	Inchoative/reflexive
50	P. se quemó _[0intent] a sí mismo _[+intent]	Reflexive only
51	P. se quemó _[0intent] intencionalmente _[+intent]	Reflexive only
52	P. se quemó _[0intent] accidentalmente _[-intent]	Inchoative only
53	P. se quemó _[0intent] accidentalmente _[-intent] *a sí mismo _[+intent]	Intent cannot be mixed
54	P. se afeitó _[0intent] en la barbería _[+intent] *a sí mismo _[+intent]	
55	P. se suicidó _[+intent] *accidentalmente _[-intent] /?a sí mismo _[+intent]	
56	P. asesinó _[+intent] a Juan (?intencionalmente _[+intent])	
57	Se rompió una pierna (para tener más vacaciones/a sí mismo) _[+intent]	=Él mismo se rompió la pierna
58	Victor se spală [PP *(pe) sine (însuși)]	It is himself that Victor is washing
59	[DP Victor (însuși)] se spală.	It is V. himself that {is washing/getting washed/washes himself}
60	*[DP Ion însuși] se spală [PP pe sine însuși]	*It is John himself that it is is washing himself
61	Ion *(se) spală pe sine	*It is himself that Victor is washing
62	*Pe sine regret că s-a murdărit Mihai	Himself I regret that Mihai got dirty

Alboiu *et al.* (2002) arrive at the same conclusions (i.e. emphatics are adjuncts not arguments) for Italian, European Portuguese, and Romanian. Notice that, in Romanian, the emphatic modifies the subject DP (59) or the noun phrase within the emphatic PP associated with the internal argument position (58), depending on whether emphasis is placed on the agent or patient, respectively. The availability of emphatics for both agent and patient further reinforces the claim that non-active SE-constructions are structurally transitive. Semantically,


emphatics mark contrastive focus (i.e. “focus logophors” *in sensu* Reinhart & Reuland 1993). Since no more than one XP can be contrastively focused in a sentence, no more than one such emphatic can appear in the argument structure of a predicate (60).

Syntactically, adverbial phrases are clearly adjuncts. In principle, PP emphatics be could arguments, however, syntactic diagnostics show that reflexive emphatics are not argumental in Romance. These emphatics fail to reflexively mark the predicate, (61) which is ungrammatical in the absence of SE, indicating that PP emphatics are SELF logophors i.e. non-argumental SELF anaphors (Reinhart & Reuland 1993). They fail numerous argument diagnostics (Hornstein 2001) e.g. extraction of SELF logophors out of factive weak islands is barred (62), confirming their adjunct status. See §3.4.4 for further discussion of reflexive emphatics.

4.2.5 Reflexives ≠ Intransitive

The above discussion assumes that clitics involved in semantic reflexivization are base-generated in argument positions receiving θ -roles i.e. they are syntactically transitive. Based upon properties shared by reflexives and unaccusatives, of inducing BE_{AUX} selection and subsequent past participle agreement in languages which show BE_{AUX}~HAVE_{AUX} distinctions such as Italian (66-67), some analysts (e.g. Grimshaw 1982 and McGinnis 2004) propose that reflexives are intransitive. In this case, SE is not a verbal argument with a θ -role, but purely a marker of a lexical process of reflexivization. Labelle (2008) notes that if reflexive verbs were intransitive (potentially involving one argument with a complex θ -role, cf. Reinhart & Siloni 2005), it should be impossible to focus more than one argument/ θ -role, but this prediction is empirically false (63). The two arguments/ θ -roles can be focused independently of each other, suggesting that the clitic carries a θ -role i.e. is an argument.

Table 64

63	Jean-Pierre s' est dénoncé lui-même	Jean-Pierre denounced himself...	French
64		(a) ...it was not others who denounced him	
65		(b) ...he did not denounce others	
66	Le ragazze Ø _{NOM} hanno fumato un sigaro	The girls have smoked a cigar	Italian
67	Le ragazze si _{NOM} sono fumate un sigaro		
68	Change of Location » Change of state » Continuation of state » Existence of state »		
French			
English			
Dutch			
Italian			
69	S' a _{HAVE} ssamuna-u i m'manuzu	S/he washed his/her hands	Làconi (Sardinian)
70	S' ε _{SBE} samuna-u/ða	S/he (got) washed	

The similarities between reflexives and unaccusatives, therefore, require a different explanation (see Alsina 1996; Doron & Rappaport Hovav 2007; Reinhart & Tal 2004; Alencar & Kelling 2005; Siloni 2008; Marelj & Reuland 2013; Sportiche 2014; *i.a.*). In fact, the two phenomena must be distinguished, because the relationship does not hold cross-linguistically. Selection of BE_{AUX}~HAVE_{AUX} does not follow the strict dichotomy proposed in the first place, whilst some languages make choices based on reflexivity~non-active.

Sorace (1992) proposes (potentially universal) continua based on aspect which progressively distinguish core unaccusative (≈“telic dynamic change”) and unergative verbs (≈“atelic non-dynamic activity”). The closer to a core a verb is, the stronger the link between its single argument and realization as internal or external argument, and the more determinate its syntactic status as unaccusative or unergative. Thus, some unaccusative/unergatives are more unaccusative/unergative than others. Languages choose different “minimal triggers” of unaccusativity along these continua (68): e.g. in Italian, “existence of state” is sufficient to guarantee unaccusative status, whereas in French (with its narrower class of syntactically unaccusative verbs), the crucial component is “change of location”. The theory predicts that

(a) the greater the distance between the minimal trigger and the core, the larger the class of syntactic unaccusatives, and the more degrees of variation a language displays, and (b) verb categories adjacent to the minimal trigger exhibit a higher degree of syntactic variation.

Cennamo & Sorace (2007)'s study of Paduan shows that inherent lexical aspect determines auxiliary choice with core verb categories, whereas compositional aspect (i.e. the event structure of the whole predicate) affects auxiliary selection with peripheral verb categories. The degree of sensitivity to these factors increases for non-core verb types as they become more distant from the core. Crucially, these choices are sensitive to age differences: in general, younger speakers tend towards the Italian model with its more extensive use of *essere*. Thus, the point of division within intransitives can change over time.

Làconi (Sardinian, Manzini & Savoia 2005) matches AUX selection to meaning: in (69), the participle displays default agreement and HAVE_{AUX} showing that SE is possessor of the *hands* ([+R,+E,DAT]=SE_{DAT}); in (70), the participle agrees with the 'patient' subject, using BE_{AUX} to indicate an internal process ([+R,-E,DAT]=SE_{MID}). Under our model, in addition to the nature of the verbal root, Italian/French are sensitive to the feature [+R] i.e. reflexive [+R,+E] and [+R,-E] non-active SE trigger BE_{AUX} and past participle agreement. Làconi Sardinian must be also sensitive to [+E], since reflexive [+R,+E] triggers the effect, but non-active [+R,-E].

Whatever the precise details, these phenomena must remain separate and cannot be used to argue for syntactically intransitive analyses of reflexives. In our opinion, this change in viewpoint is not a loss of an important semantic/syntactic insight, but the correction of an empirically unfounded over-generalisation.

4.2.6 Anticausatives ≠ Reflexives

Some authors (e.g. Chierchia 2004 for Italian; Koontz-Garboden 2007, 2009 for Spanish; also Beavers & Koontz-Garboden 2013a, 2013b) expressly define anticausativization and reflexivization as the same process, characterising such clauses as ‘The glass caused its own breaking’. Piñón 2001, Doron 2003, and Folli 2003 *i.a.* question whether anticausatives such as ‘the boat sank’ really mean ‘(some property of) the boat sank the boat’; even more so, cases like ‘The wound healed within two weeks’. Doron (2003) mentions Hebrew’s anticausative version of ‘give birth’ where reflexive interpretations are inconceivable (X gave birth to X). There have been numerous rebuttals on technical grounds (e.g. Horvath & Siloni 2011, 2013; Alexiadou *et al.* 2015).

Despite identity in morphological shape, the two classes (71-73)~(74-76), differ semantically in their adicity. Only SE-reflexive verbs are semantically transitive predicates with external and internal θ -role, which are both assigned to the same entity via binding of internal argument by external argument (73). SE-anticausatives are semantically intransitive predicates with an internal θ -role only (76). This can be shown by the fact that the transitive counterpart of SE-anticausatives logically entails the SE-anticausative (i.e. (74) entails (75), that the glass is broken), while the transitive counterpart of SE-reflexives do not entail the SE-reflexive verb ((71) does not entail (72), that John washed himself, but rather than John was/became washed). The SE-morpheme works as a (locally) bound variable only in SE-reflexives.⁵⁴ Crucially, both SE- (79) and unmarked (81) anticausatives can take reflexive readings, which would be impossible if reflexivity=anticausativity (Schäfer & Vivanco 2015).

⁵⁴ Doron & Rappaport Hovav (2007), Spathas (2010), and Sportiche (2014) for tests showing that SE-reflexives should be analyzed as bound variables and not as reflexivizers.

Table 65

71	Sa mère a lavé Jean	(transitive)	His mother washed John
72	Jean s' est lave	(SE-reflexive)	John washes himself
73	[[se laver]] = $\lambda x \lambda e$ [wash(e) \wedge AGENT(e, x) \wedge PATIENT(e, x)]		
74	Jean a cassé le verre	(transitive/causative)	John broke the glass
75	Le verre s' est cassé	(SE-anticausative)	The glass broke
76	[[se casser]] = $\lambda x \lambda s \lambda e$ [BECOME(e, s) \wedge broken(s) \wedge THEME(s, x)]		
77	John does not have four children . He has three dogs	Propositional Negation	
78	John does not have four children. He has five children	Metalinguistic Negation	
79	El vaso no se rompió a sí mismo, pero tú lo rompiste	The glass didn't break itself, but you broke it	
80	Los precios aumentaron	The prices increased	
81	Los precios no se aumentaron a sí mismos, pero A. los aumentó	The prices didn't increase themselves, but A. increased them	
82	Las tostadas se quemaron	The toasts { ^v got burned/#burned themselves}	
83	La puerta se abrió	The door { ^v opened/#opened itself}	
84	La puerta automático se cerró	The automatic door { ^v closed/ ^v closed itself}	
85	El niño se quemó	The kid got { ^v burnt/ ^v burnt itself}	
86	The vase broke by itself		
87	John broke the vase by himself		
88	*The vase was broken by itself		
89	Maria ha dovuto suggerire la risposta? No, Gianni sapeva la risposta da sé Did Mary have to suggest the answer? No, Gianni knew the answer by himself		
90	Non devi asciugarli. Diventeranno asciutti da sé You do not have to dry the dishes with a towel. They become dry by themselves		
91	Non innervosire Maria! Diventa già' nervosa da sé! Do not make Mary nervous! She gets nervous already by herself		

Koontz-Garboden (2009) argues that negation proves that entailment is not maintained (79). Such examples, however, involve ‘metalinguistic’ negation (e.g. Horn 1985) not negating the truth-value of the proposition (77) but objecting to some pragmatic aspect of it (78). When speakers do not want to (or cannot) identify cause, anticausative expressions are more appropriate (*in sensu* Higginbotham 1997) than corresponding (active or passive) causative constructions (cf. Rappaport Hovav 2014). Appropriateness depends upon perspective; if the hearer disagrees with this choice, anticausative verbs may be metalinguistically negated in order to object to and modify the scalar implicature (e.g. four=‘four and no more’) associated

with the verb. Negation in (79) does not deny that the vase is broken, but objects to the implicit denial of responsibility created by speaker selection of the anticausative construction.

Reflexive readings with inanimate subjects are generally avoided: ‘The glass broke itself’ does not convey lack of identifiable cause, but inappropriate personification, hence the default reading is one of anticausativization. When that construal is negated and/or enforced by intensifiers, reflexive readings become available (79), just as unmarked anticausatives may also (80) take reflexives under such circumstances (81). Depending on context, intensifiers are sometimes not even necessary. Whilst most inanimate objects do not act under their own volition (82-83), an automatic door, designed to close itself would be acceptable with a reflexive reading (84) without *a sí mismo*. This holds even more strongly for human DPs, because they are capable of more actions (85), where the default reading is reversed, since sentient beings don’t tend to wilfully damage themselves (§4.2.4, masochistic reading). The interplay in (85) is between middle/passive and reflexive.

Contra Koontz-Garboden (2009),⁵⁵ far from adding masochistic sentience, *by itself* reinforces the entailment, meaning ‘unaided’, rather than ‘through its own activity’ (86). By its use, speakers *assert* the lack of causer i.e. nothing can be identified as causing John to break the vase (87), or the breaking event itself (86). In (86), *by itself* stresses that use of causative constructions (i.e. involving an external cause(r) argument) is unjustified on the basis of their knowledge.⁵⁶ Contra Koontz-Garboden, *by itself* does not identify, but rejects the participation of a causer. In (88), *by itself* makes the same semantic contribution as in (87), but creates a contradiction exactly because periphrastic-passive semantics contain an implicit causer.

⁵⁵ Also Chierchia (2004:42) for Italian *da sé*.

⁵⁶ For similar conclusions regarding *da sé*’s counterparts in other languages, see e.g. Reinhart 2000, Pytkäinen 2002, 2008; Alexiadou *et al.* 2006a, 2006b.

By definition, these verbs are associated with their lexical causative alternate, regardless of SE-marking. Across languages both types license *by itself* (Alexiadou *et al.* 2006a; Schäfer 2008a; Horvath & Sioni 2013). With predicates lacking a causative counterpart, *by itself* is difficult to use. This is the case with existing transitives (89), and inchoative structures that lack a lexical causative counterpart e.g. pure unaccusative verbs like *blossom* or eventitive copula constructions (90, 91). However, once it is contextually established that the events expressed could, in principle, be caused, then exclusion of such causation through *by itself* becomes available *in order to deny that possibility* (Horvath & Sioni 2013:220; more examples in Alexiadou *et al.* 2015). Thus *por sí mismo* performs a similar task to *a sí mismo* in intensifying the existing statement.

Nothing (except world knowledge) blocks semantic reflexivization of causative verbs, although reflexive readings tend to require contextual support. Given that verbs can operate both reflexively and non-actively (i.e. as passives, middles, and anticausatives), the two concepts/uses must be kept separate. Conversely, we cannot impose overly complex methods of attaining either, whereby one method denies the other. This is a problem for many approaches which are designed to eliminate the possibility of one or more of (77-91).

Reflexivization approaches cannot accommodate the fact that SE-reflexives, but never SELF-reflexives, are used across languages to mark anticausatives (Faltz 1985; Kemmer 1993; *i.a.*), because the semantic outcome of the two reflexivization strategies is identical. SE-marking of non-active and reflexive constructions found across languages is a real syncretism (same form, different function) as illustrated in §2.2.1-2.2.2.

4.2.7 Conclusions for Reflexivity

Reflexive/reciprocal clauses are open, near-reflexive constructions, with *semantically* distinct arguments. Reflexive/reciprocal *se* does not reduce the predicate's *semantic* valency, but classifies predicates as reflexive by filling argument positions with reflexive forms. The slot filled by *se* (i.e. its case) determines the 'view' being taken of the subject; theme, patient, or agent. In the remainder of this chapter, we will argue for a matching set of non-active uses.

4.3 Non-Reflexive SE

Middle voice is traditionally seen as showing “the action is performed with special reference to the subject” (Smyth 1920:§1713) or the subject “inside the process of which he is the agent” (Benveniste 1966b:149). Cross-linguistic data (Kemmer 1993, 1994; Maldonado 1988, 1992, 1993, 1999; *i.a.*) suggest that rather than focus on agents, middle constructions highlight changes-of-state experienced by grammatical subjects within events. While transitive active constructions depict situations of two (possibly coreferent) participants interacting, middle voice involves only the subject/experiencer, where (unlike reflexives) it is impossible to distinguish separate images of that participant, and thus contrast between them (as introduced by *lui-même/sí mismo*) is meaningless.

Table 66

	a) Reflexive	b) Middle
92	Meg-üt-ött-e mag-á-t PERF-hit-PAST-3SG SELF-his-ACC He hit himself	Bele-üt-köz-ött- (valami-be) PERF-hit-self-PAST-3SG.INDEF (something-ILL.) He bumped into something
93	On utixomiril sebja He pacified REFL He controlled himself.	On utixomiri+sja He pacified+MID He calmed down.
94	Se paró a sí mismo He stood himself up (paralysed patient)	Se paró He stood up

Many languages display reflexive~middle distinctions in surface-form (Haiman 1983:797). In

two-form languages, middles and reflexives are expressed by different markers, where reflexives are normally longer, e.g. Hungarian (92, reflexive pronoun *magat* vs. verbal suffix *-kod/-koz-*) or Russian (93, reflexive pronoun *sebja* vs. verbal suffix *-sja*). Haiman (1983) considers this “iconic”; complexity reflecting degree of event elaboration, where long-form reflexives express split-representations, short-form middles highlight single representations. Whilst Romance has single-form languages using SE for both constructs, distinction is achieved by adding *lui-même/sí mismo*, making reflexives longer. In (94), *sí mismo* produces a reflexive reading where the subject acts upon their body as if it were a separate object as opposed to the single internal event of ‘standing up’.

Table 67

Inalienable possession	≈grooming or body care	<i>Lavarse</i> ‘wash’, <i>peinarse</i> ‘comb’
Self-benefit actions	≈benefactive middle	<i>Conseguirse</i> ‘get’, <i>allegarse</i> ‘obtain’
Non-translational motion	≈change in body posture	<i>Pararse</i> ‘stand up’, <i>sentarse</i> ‘sit down’
Internal change (emotional)	≈emotional reaction middle	<i>Alegrarse</i> ‘gladden’, <i>enojarse</i> ‘anger’
Manifestations of emotions	≈emotive speech actions	<i>Quejarse</i> ‘complain’, <i>lamentarse</i> ‘lament’
Internal change (mental)	≈cognition middle	<i>Acordarse</i> ‘remember’, <i>imaginarse</i> ‘imagine’
Change in location	≈translational motion	<i>Irse</i> ‘leave’, <i>subirse</i> ‘ascend’
Changes of state	≈spontaneous events	<i>Romperse</i> ‘break’, <i>abrirse</i> ‘open’

Cross-linguistically, there are consistent situations which lend themselves to middle encoding (Kemmer 1993, 1994), illustrated for Spanish in Table 67 (Maldonado 2008). This is a heterogeneous set with agentive or patient properties; some are punctual/inceptive, others are durative/inchoative. These represent distinct categories and correspond to different marking schemes in other languages, e.g. in English such cases may be expressed by intransitive verbs marked by *-en* (‘sadden’); extended by particles (‘stand up’); or by aspectual verbs (‘got sick’). There is clearly a need for further division of Romance non-active SE.

Our approach is motivated by three key considerations. In §4.3.1, we argue against approaches that convert one underlying form into another and for a common base approach, allowing languages to use different, often multiple, means to mark different subsets of non-active constructions. In §4.3.2, we extend this approach by arguing that one of the methods of marking is often the lexical label itself. §4.3.2 also argues for the acceptance of more patterns of usage than ordinarily taken into account, which we consider reflect underlying structures (even if not as overtly as the languages discussed above). In our proposal in §4.3.4, each pattern reflects a different kind of non-active SE defined in terms of case. In §4.3.3 we argue that applicability, and the possibility of interpretation, of such patterns is defined entirely by the properties of the overt ‘patient’ subject within the construction (within context). We reject complex connections to implicit arguments as inconsistent.

These considerations lead to a very different and much simpler proposal (§4.3.4) than usually found in the literature, but one which is able to reflect patterns of usage discussed in §4.3.2 and builds on the theoretical approach taken for reflexives in the previous sections. We believe that this is fundamentally necessary, since many uses are not only vague between non-active types but across the non-active~reflexive divide. A concept which we return to at the close of the chapter.

4.3.1 Morphological Marking

The most common division of non-active forms is between passives (discussed at length in §4.6) and anticausatives, the latter being the product of a ‘causative alternation’. Cross-linguistically, such alternations show wide variation in morphological marking:⁵⁷ Polish marks

⁵⁷ Haspelmath (1993) for an overview, and Piñón (2001); Doron (2003) for further discussion.

anticausatives (95), Khalka Mongolian marks causative variants (96), Japanese derives both variants from a common stem (97), whilst English shows no distinction (98). Often languages display different roots to represent causative~anticausative (99). For languages which mark anticausative variants, some mark all anticausatives (e.g. Polish), whilst others mark only a subset (e.g. English). Many of the latter languages possess verbs which occur in both arrangements, and verbs which cannot enter into the alternation (100-101). In Romance, some anticausatives must occur with SE (102), others remain necessarily unmarked (103), whilst some (104-105) are optionally marked (Centineo 1995).

Table 68

	Intransitive	Transitive		
95	złamać- się	złamać	‘break’	Polish
96	ongoj-x	ongoj- lg -ox	‘open’	Khalka Mongolian
97	atum- aru	atum- eru	‘gather’	Japanese
98	break	break		English
99	die	kill		
100	bloom	x		
101	x	murder		
102	La finestra *(si) è chiusa		The window closed	Italian
103	La temperatura *(si) è diminuita		The temperature decreased	
104	Il cioccolato è fuso per pochi secondi/in pochi secondi		The chocolate melted for/in a few seconds	
105	Il cioccolato si è fuso *per pochi secondi/in pochi secondi			

Haspelmath’s (1993) typological study proposes a universal ranking of predicates along a “spontaneity scale”. If languages morphologically mark a particular transitive verb, they will also mark all other transitives expressing events of equal/higher spontaneity i.e. implied lack of agentivity. Correspondingly, if languages mark an intransitive verb, they will mark all other intransitives expressing events of equal/lower spontaneity.

Theoretical approaches may be divided on the basis of their starting conditions. Many start from intransitive (anticausative/unaccusative) entries which are converted into transitives

through a process⁵⁸ of *causativization* and marked by the extra morphology. However, such an approach cannot explain languages which mark (a subset of) their anticausative alternants, as these are assumed to be basic/underived (Harley 1995 and Folli 2003 for various proposals). Moreover, non-alternating verbs must be further restricted by the verb's lexical entry. Thus, even Ramchand (2008)'s syntactic approach requires a lexical component to determine that English *murder* obligatorily occurs in transitive/causativized syntactic structures. The opposite direction of derivation (i.e. starting from transitive bases) has also been proposed within lexicalist theories,⁵⁹ but face the reverse problem regarding morphology. Since anticausatives are assumed to be derived from causative variants, morphology found on anticausatives can be seen as marking a derivational process, but languages that mark (a subset of) their causative alternants cannot be accounted for.

One might posit that both processes exist across, or even within, languages e.g. Brousseau & Ritter (1991) for French. Alternatively both variants may be derived from a common base. Languages may differ in whether they mark one, none, or both derivational processes. For lexicalist approaches, see Davis & Demirdache (2000) and Piñón (2001), and syntactic approaches, Pylkkänen (2002, 2008), Embick (2004a, b) or Alexiadou *et al.* (2006). This work follows the common base approach. Not only because it appears to be the only practical solution, but also because it provides a more direct link between construal and surface form.

4.3.2 Variations

Many verbs operate transitively and intransitively. For *terminar(se)*, intransitive achievement readings (–SE) apply to situations with (106a) and without (108a) plausible external cause.

58 Lexical (e.g. Hale & Keyser 1986, 1987), or syntactic (e.g. Harley 1995; Pesetsky 1995; Folli 2003; Folli & Harley 2005; Ramchand 2008).

59 e.g. Grimshaw (1982), Chierchia (2004), Levin & Rappaport Hovav (1994, 1995) and Reinhart (2000/2002).

With SE, readings are inceptive (107a/109a), focusing on the pivotal moment of change. As a transitive, *terminar(se)* displays causation (106b) and its passivization (107b). Whilst (110a, –SE) implies that the *fiesta* came to a natural end (without external cause) and might receive (111a) as an answer, (110b, +SE) must be read with implied external force, and appropriately answered (111b). It is the passive equivalent of the causative transitive (106b), i.e. the mass *was terminated*. The (b) variations are impossible with *verano* (108/109). Acceptability is determined by the agent’s ‘teleological capabilities’: winter can stop, but no agent may stop it. Intransitive-SE is *not* indicative of causative-alternation: (107b) is not the ‘alternate’ of (106a), but (106b). (107a) is the ‘alternate’ of equally non-causative intransitive (106a). Similarly, (109a) to (108a), which do not possess causative variants (108b-109b).

For *dormir(se)*, intransitive activity (112a) alternates with intransitive inceptive (113a); neither implying external cause. *Dormir* may also operate as causative transitive (112b),⁶⁰ but since the verb describes an ongoing state (not bounded achievement), it cannot (unlike *terminar*) alternate with de-causative or passive (113b, +SE). *Aparecer/crecer/morir* describe changes-of-state of subject inherent properties beginning and ending within their subject, without need for external cause. Both *aparecerse* (119a) and *crecerse* (121a) are restricted to [+ANIM] entities capable of intentionally changing. *Morir* requires animacy by virtue of its meaning, but is available in [±cause] contexts. *Morirse* highlights ingression into the new state of death, but does not include intention. It may only be used in contexts in which it is understood as a natural process without external cause (Otero 1999:1467).

⁶⁰ Transitive use is rare in Spanish, but occurs: *El gentil monstruo durmió a su amigo a punta de caricias* (CREA). French shows much wider use of ‘inherent reflexives’ without SE e.g. *Je couche les enfants à 20h*, ‘I put the children to bed at 8pm’.

If *morir(se)* were a causative alternation, (115a) would be equivalent to (115b). While *morir* allows non-human causes expressed peripherally (117a), they become awkward with *morirse*: whilst grammatical, it is far less acceptable/usable with a distinct meaning (*≈la explosión inició un proceso que provocó la muerte natural de Pedro*). (117b)~(117a) differ solely in role assignment; *explosión* as actor (117a) vs. secondary effectuator (117b). *Morir(se)*~*matar(se)* is equivalent of *crecer(se)*~*cultivar(se)*, *aparecer(se)*~*demostrar(se)*, construals of similar concepts using different lexemes to which have aggregated different subtleties of semantics which make them close but not interchangeable.

Columns a~b of Table 69 represent *two* verbs each with the possibility of interacting with SE. In some cases, the same surface-form is used to express both columns (e.g. *terminar*), in others a different form is selected (e.g. *morir*~*matar*). This kind of lexicalization is historically quite flexible. French transitive *tuer* ‘kill’ has developed *se tuer* ‘die’ e.g. *Il s’est tué dans un accident*, ‘He died in an accident’, where self-affectedness interpretations (unintentional death) require contexts excluding agentivity (*dans un accident*), otherwise the construction is read as suicide. Conversely, the Spanish Rio Platense dialect, regularly ‘transitivizes’ unaccusatives e.g. *Juan no murió, lo murieron* ‘John didn’t die, they killed (lit. died) him’ (Pujalte & Zdrojewski 2013). We conclude that the main reason that verbs do not enter the ‘causative alternation’ is that the same concept already has a surface form of its own.

Table 69

	(a) Intransitive		(b) Transitive	
106	La misa terminó The mass ended	Achievement	El sacerdote terminó la misa The priest ended the mass	Causative Achievement
107	La misa se terminó The mass came to an end	Inceptive	La misa se terminó cuando... The mass was terminated when...	Passive
108	El verano terminó The winter ended	Achievement		
109	El verano se terminó The winter came to an end	Inceptive		
110	¿Cómo terminó la fiesta anoche? How did the party end?		¿Cómo se terminó la fiesta anoche? How was the party brought to an end?	
111	Bien; todos nos fuimos muy contentos Well, we all left very contented		Llegó la policía y todos escapamos The police arrived and we all fled	
112	El niño durmió 5 The child slept	Activity	La madre durmió al niño The mother got the child off to sleep	Causative
113	El niño se durmió 6 The child fell asleep	Inceptive	*El niño se durmió *The child was slept	
114	Pedro murió Pedro died	Achievement	Pedro mató a Juan Pedro killed Juan	Realization
115	Pedro se murió Pedro died	Inceptive	Pedro se mata Pedro was killed	Passive
116			Pedro se mató (a sí mismo) Pedro killed himself	Reflexive
117	Pedro (?se) murió con la explosión =la explosión mató a Pedro		Pedro se mató con la explosión =la explosión mató a Pedro	
118	√La tinta invisible/√Pedro apareció The invisible ink/Pedro appeared	Realization	demonstrar	
119	*La tinta invisible/√la Virgen se apareció The Virgin allowed herself to be seen	Inceptive	demonstrarse	
120	√El manzano/√El atleta creció The apple/athlete grew	Realization	cultivar	
121	*El manzano/√El atleta se creció The athlete grew (in stature/skill)	Inceptive	cultivarse	
122	La lluvia cae The rain is falling	Activity	✗	
123	Adrián se cayó Adrian fell down	Inceptive		
	[–CAUSE]		[+CAUSE]	

Rather than explain this array of data in terms of transformations of one form to another, we start from a semantic concept (e.g. morbidity), and by selecting the construal for its use (e.g. entering into death, being caused to enter into death, being dead), access the associated morphological/phonological label ('die', 'kill', 'be dead'). The fact that some languages have

the same label under more than one heading is a matter of historical accident, not evidence of syntactic process. Rather than an alternation between transitive+Ø~intransitive+SE, there is a range of surface forms based on intransitives±SE and transitive±SE, which sometimes have the same label.

Some verbs do not have a matching transitive concept (and, therefore, no label to express it). Verbs such as *caer* cannot enter into the causative alternation (i.e. unlike (118-121) there is no meaningful counterpart to ‘fall’), yet they can alternate between –SE and +SE (122-123). This is the same alternation seen with *morir(se)*_{INTRANS} (which has *matar(se)*_{TRANS} as its counterpart) and with *terminar(se)*_{INTRANS} (which has *terminar(se)*_{TRANS} as its counterpart). Thus, application of SE to intransitives is quite distinct from its application to transitives. Moreover, SE has more than one function with transitives, representing middle and passive, as well as anticausative readings; the latter case bringing it into line with the intransitive+SE cases.

4.3.3 Restrictions on Application

Which roots can undergo the ‘causative alternation’ varies across languages. (124) is acceptable in Hebrew, but not English. (125) is acceptable in English, but not French/German (Levin & Rappaport Hovav 1995:113ff). McKoon & Macfarland (2000) and Wright (2002) show that many unaccusatives presumed by Levin & Rappaport Hovav (1995) to lack causative variants can in fact be found in corpora in causative uses. Often, it is context which determines acceptability (126). For intransitive verbs without transitive counterparts (e.g. bloom, blossom), Chierchia (2004) and Reinhart (2000, 2002) claim that most have transitive counterparts in some language, hence their absence is simply a lexical gap. As we have seen, in many cases, the lexical gap is in fact filled by another surface-form.

Table 70

124	He danced	~The musician danced him (i.e. made him dance)
125	The bicycle leaned against the fence	~I leaned the bicycles against the fence
126	The presenter danced her right off the stage	

The question of what determines the possibility of alternance has been hotly debated (Hale & Keyser 1986; Levin & Rappaport Hovav 1995; Reinhart 2000, 2002; Härtl (2003); Alexiadou et al. 2006A, 2006b; i.a.). Levin & Rappaport Hovav (1995) conclude that causatives that restrict their external argument to agents (or agents and instruments) and disallow causer cannot form anticausatives. Reinhart (2000, 2002) states that only those causatives that leave the nature of their external argument unspecified form anticausatives (127-130). Whilst these exemplify a cross-linguistic generalization, it is not perfect. Some languages (e.g. German, Härtl 2003) have a small class of alternating verbs which restrict their external argument to causers and exclude agents, contra Reinhart (2000, 2002). Other languages have a larger group of verbs with unrestricted subject but, nevertheless, do not form anticausatives e.g. English ‘kill’ and ‘destroy’. Some languages restrict their external arguments to agents and never license causers e.g. Jacalteco (Craig 1976) and Japanese (Yamaguchi 1998). Nevertheless, these languages have the ‘causative alternation’.

Table 71

127

The vandals/The rocks/The storm broke the window	The window broke
128 John/the hammer/storm enlarged the hole in the roof	The hole in the roof enlarged
129 The terrorist/*explosion murdered the senator	The *senator/*explosion murdered
130 John/*wind removed the sand from the rocks	*The sand removed (from the rocks)
131 He broke his promise/the contract/the world record	*His promise/The contract/The world record broke
132 The bad weather broke	

133

Anticausative	Middle	Passive	Transitive/Active
The vase broke	Vases break easily	The vase was broken	He broke the vase
*The promise broke	Promises break easily	The promise was broken	He broke the promise
The bad weather broke	*Bad weather breaks easily	*The bad weather was broken	*He broke the weather

Alternations are restricted by the nature of their central participant. Intransitive *break* (133) shows stronger selectional restrictions on its theme than transitive *break* (Levin & Rappaport Hovav 1995 for more examples). Levin & Rappaport Hovav explain such cases through world-knowledge; (131) necessarily involves an intentional agent, but this does not follow since (132) is perfectly acceptable. Levin & Rappaport Hovav (1995:107) argue that “what characterizes the class of alternating verbs is a complete lack of specification of the causing event”, which is reflected in the wide variety of subjects admitted by these verbs. Whilst we accept the first argument, there is little connection between this statement and the subjects of their causative variant. What matters is the anticausative subject’s “teleological capabilities” i.e. bad weather is capable of breaking without external causation, promises/contracts are not. The agency of some putative external cause(r) is irrelevant. Objects such as vases cross the border. They are entities requiring external agents in order to break (hence capable of taking passive readings), but they can also be seen as items that sometimes ‘just break’ leading ‘accidental causer’ readings. In §4.4, we show that non-active constructions do not define any external agent. They can only (sometimes) be inferred from context, and sometimes implied as a reading on OBL as an ‘accidental causer’. From our perspective, the limitations on application of any particular meaning of SE are determined solely by the capabilities of the *single* participant subject undergoing the event.

As with morphological classes discussed in §2.2, there is a tendency towards reductionism. Analyses seek simple answers where, what are from our point of view, distinct items are subsumed under generalisations based on a limited number of similarities. Our argument is that not only do we need to start from a common base, but we need to recognize more targets

i.e. surface patterns indicating underlying structures. In order to understand non-active SE, we need to recognize all its uses, be able to contrast it with its distinct active reflexive uses and the distinction between its presence vs. absence with intransitives e.g. ‘the vase broke (–SE)’ and ‘the vase broke (+SE)’. Without understanding the full range, we cannot understand non-active SE’s place in the overall framework.

Any attempt to subsume one construction under another will inevitably lead to error. We believe this is the wrong approach. We should accept the vagueness of surface forms, identify the distinct units, explain how they came about, and explain how such similar forms can be interpreted. In line with the programme set out in the introduction of this work, this chapter does not provide a detailed syntactic account of these phenomena, but rather classifies real (not idealised) usage and provides a single coherent model to explain not only the range and distinctions, but also how the surface overlap can be interpreted (i.e. easily distinguished in context) in order to perform its communicative function.

4.3.4 Proposal

‘Closed’ predicates (§4.2.1) possess unique variables saturating two thematic roles. Pure-reflexives are used for true identity. Just as reflexive/reciprocal SE (‘open’ predicates) takes the case appropriate to the relationship between participant and verb, non-active SE (‘closed’ predicates) also appear in different cases.

Reflexive constructions require “conceptual separation” (Kemmer 1994:206-9) of one entity into two distinguishable roles: actor vs. external self as effector, affectee, or effectee. Non-active SE focuses on subject-internal events as seen from the perspective of those same roles, making other arguments irrelevant and demoting cause to circumstantial expressions: *de-causatives* (Geniušiene 1987:319-24). Inanimate subjects (unless personified) eschew reflexives since they cannot create ‘conceptual separation’, but often appear in middle constructions describing an internal COS. In this model, ‘middles’ are subcategorized by case i.e. the secondary role being highlighted: theme (accusative/passive), patient (dative/middle), agent (nominative/anticausative). In order to avoid confusion with terminology found elsewhere, the remainder of the document restricts itself to use of SE_{ANT}/SE_{MID}/SE_{PASS} as defined here (Table 72), and SE_{IMP}/SE_{NAR} are set out in §4.6.⁶¹

Table 72

[±R]	[±E]	NOM	DAT	ACC	Morphology	$f(x)$
+	–	SE _{ANT}	SE _{MID}	SE _{PASS}	Non-active	Self
	+	SE _{NOM}	SE _{DAT}	SE _{ACC}	Reflexive/reciprocal	Self by reference
–	–	Agent	Indirect-Object	Direct-Object		Other
	+	Instrument	Locative	Partitive		Other by reference

⁶¹ The nature of [–R,+E] entities is developed in Chapter 5, and spurious *se* (SE_{SPUR}) in Chapter 6.

Since SE generally enforces BE_{AUX}, there is little *surface* difference between reflexive and non-active uses. Some languages differentiate reflexive~non-active SE by virtue of auxiliary selection (§4.2.5), whilst Vernacular Brazilian Portuguese shows SE for reflexives and Ø for non-active constructions (§2.2.2), but this is rare. There are, however, notable differences. In non-active constructions, subjects (indicated by verb agreement) often remain in S_L, thereby defocusing them. SE_{PASS} does not license DAT (134-135), SE_{MID} does not accept ACC (136), SE_{ANT} voids both DAT/ACC (137). Since these restrictions affect complements as well as clitics, they cannot reflect clitic~clitic exclusions.

Table 73

Table 73

	S _H	N	O	D	A		Reading	
134		Ø _i			se _{PASS}	venden _i los libros _i aquí	Books _i are sold _i here	Passive
135					se _{PASS}	vende _i la casa _i	The house _i was sold _i	Passive
136			(le) _j	se _{MID}		abrió _i la puerta _i (a María) _j	The door _i opened _i on Maria _j	Middle
137		se _{ANT}				rompieron _i los platos _i (a él) _j	The plates _i broke _i (on him) _j	Anticausative

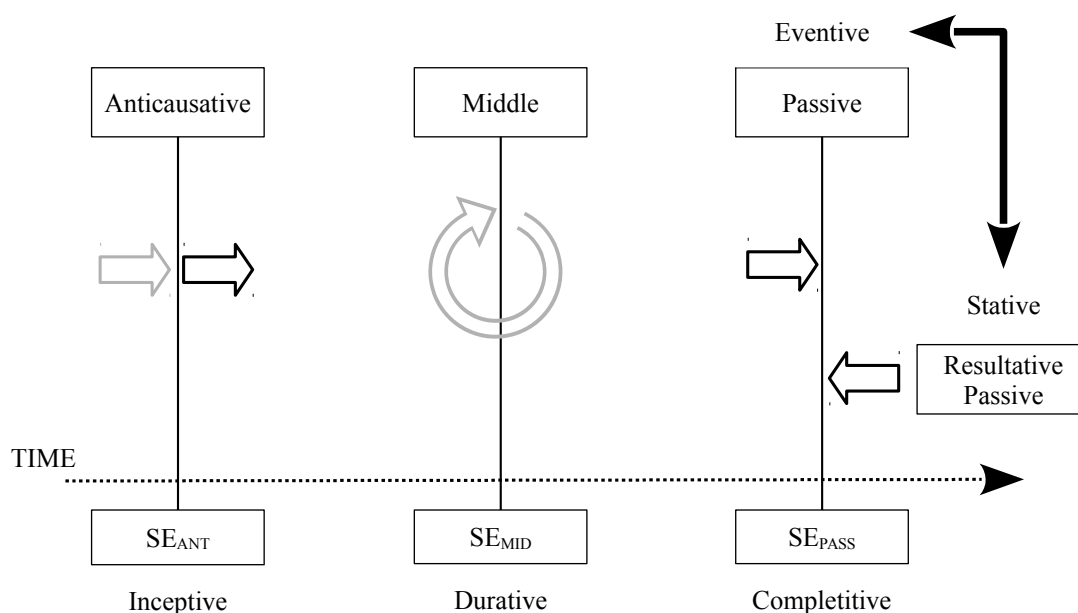
Each type of non-active SE has its own position. OBL confirms SE_{ANT} as NOM, and SE_{MID/PASS} as VP argument referents. Since passives do not license datives, and middles do not license accusatives, it follows that SE_{MID}=DAT and SE_{PASS}=ACC. The reflexive~non-active pairs (SE_{ANT}~SE_{NOM}/SE_{MID}~SE_{DAT}/SE_{PASS}~SE_{ACC}) are distinguished by focus upon transfer of energy from subject to secondary self-image in each case-defined role (active/transitive), or upon change-of-state within, and viewed from the perspective of, that role (non-active/non-transitive). In addition, OBL may display (un)intentional facilitator readings in non-active constructions unavailable with reflexives (§3.3.6).

The fact that SE_{NAR} uses the same form in Spanish (§4.6.5) but not other languages reflects phonological syncretism. Development of SE_{IMP} in many, but not all, languages is a matter of

functional syncretism. SE_{SPUR} in Spanish, but few other languages, reflects the interplay of both processes which is highly particular to this language (§6.2).

4.3.5 Properties

Non-active constructions describe the attribution of properties to their subject, for which the subject has innate potential i.e. “teleological capabilities”. Under this definition, non-active subjects are agents (although not dynamic) and so may raise to S_H or remain in S_L . Here we give a brief description of the properties of each non-active SE, which are developed in subsequent sections.



SE_{ANT}	Inceptive changes-of-state, acknowledging the struggle prior to achievement (§4.7.3).
SE_{MID}	Inchoative, ‘becoming X’, but not reaching full change-of-state. These are processes or iterations of events forming an overall process. The latter option making SE_{MID} particularly suited to generic statements e.g. books sell easily.
SE_{PASS}	A punctual event, where the subject has ‘been Xed’, prior states are irrelevant. §4.6.2 makes a clear distinction between stative periphrastic-passives and eventive SE_{PASS} .

A further distinctive property of this model is the recognition of full paradigms for each type (with some language-specific limitations on passives in Spanish). Many theories start from the basis that non-active SE is limited to 3-person. For example, Sánchez López (2002:138) proposes that non-active constructions with SE impede the presence of subjects with features for 1/2-persons, although the means or reasons are unspecified. This has the effect of splitting coherent sets e.g. animate~inanimate middles and anticausatives, whilst coalescing disparate functions e.g. passives~impersonals, leading to convoluted justifications, incoherent rule sets and amorphous operators such as transpositors/intransitivizers. The significant variable to be considered is animacy. The fact that inanimates are usually 3-person underpins the misunderstanding.

SE_{MID}: Verbs available to animates and inanimates differ because middles describe attribution of properties natural to their subject which differ based on animacy: doors do not anger, people do not open. However, if we give objects human qualities (personification), they may take on these attributes; doors can get angry. Equally, they may continue to use their existing properties in 1/2-contexts; *‘I read easily’, said the book* is no different from *‘I bribe easily’, said the politician*. Moreover, animacy represents a continuum of sentience; *gorillas get killed* but *mosquitoes are killed* (§4.8). There is no justification to separate ‘inanimate middles’ from ‘personal inchoatives’. They do not require separate classification or syntax.

SE_{ANT}: Whilst ‘middles’ imply but obfuscate external causes (*de-causative*), *anticausatives* deny them. *Morirse*+person is no more or less anticausative than *romperse*+object. The verbal attribution is simply more or less appropriate: animates live/die, but do not break; inanimates

break, but do not die. However, when personified inanimate objects show the same pattern as middles; cars can die. The degree of sentience attributed to the subject determines availability; robots do more than bricks. There is, therefore, no justification to separate out the ‘inanimate anticausatives’ from ‘personal aspectual’ uses.

SE_{PASS}: Although always limited in their use with local persons in Spanish, such passives were possible in Old Spanish, and remain acceptable in other Romance languages. The lack of Spanish personal passives is due to the overlaying of language-specific rules against this particular application, which developed hand-in-hand with the specialized SE_{NAR}+OBL construction (§4.6.6) and the development of personal-*a*. There is no justification for distinguishing 3-person inanimate passives as syntactically special at the theoretical level.

Table 74

	EFFECTOR	AFFECTEE	EFFECTEE	
– ANIM	Se _{ANT} rompieron los platos	La puerta se _{MID} abrió	Se _{PASS} venden los libros	–E
+ ANIM	Se _{ANT} murieron los cristianos	El político se _{MID} soborna fácilmente	Se _{PASS} mataron los cristianos ⁶²	–E
	Se _{NOM} comió la torta	El político se _{DAT} pagó mucho	Pablo se _{ACC} mató	+E
	Se _{NAR} <les> murió a _{OBL} los cristianos		} [–SPEC]	
	Se _{IMP} <los> murió a _{ACC} los cristianos			

Each non-active usage focuses upon the subject playing a particular role: SE_{PASS}↔effectee; SE_{MID}↔affectee; SE_{ANT}↔effector. Reflexive SE references secondary images of the subject, showing energy input into the event returning to them under the same roles. Lacking mental force, inanimates cannot project into the world, and hence cannot take reflexives. As confirmed by the fact that when personified, they can. Similar arguments based on personification are found in García Negroni (2002, Italian, and Zribi-Hertz (1982, French).

62 Old Spanish only.

4.3.6 Outline

In order to accept this arrangement, it is necessary to show that, contra many analyses, SE_{ANT}/SE_{MID}/SE_{PASS} form a logical class *without* distinctions with regard to external causer syntax (§4.4), and *with* a common underlying structure (§4.5). As part of this process, it is necessary to separate out SE_{IMP} and SE_{NAR} often confused with SE_{PASS} (§4.6), and distinguish between the two nominative uses of SE i.e. reflexive SE_{NOM} vs. non-active SE_{ANT} (§4.7). In (§4.8), we show that the proposed model of non-active SE has the ability to explain the full range of uses found in real language (not idealised descriptions) whilst leaving little room for miscommunication.

4.4 Non-Actives as a Class

The standard view is that active subjects are suppressed (but still accessible) in passives but deleted in marked anticausatives. Presence vs. absence is determined by diagnostics including, control into purpose/adjunct clauses, instrumental/agentive adverbs, availability of *by*-phrases (Manzini 1983; Roeper 1987; Baker *et al.* 1989; Levin & Rappaport Hovav 1995; Reinhart 2000; *i.a.*). In our opinion, such tests do not carry the significance afforded them.

Contra the commonly held view that verbal passives semantically always include an implicit argument (e.g. Bhatt & Pancheva 2006 and references therein), we argue for a distinction between periphrastic- and SE-passives. Whilst the former may, Romance SE-passives, like SE-middles and SE-anticausatives do *not* contain such implicit arguments. All non-active SE-constructions align with unaccusatives in only possessing a single ‘patient’ argument.

4.4.1 Event-Passives

Event-passives⁶³ (e.g. Solstad 2007a, 2007b) are verbal passives which involve only caused events with no volitional agent present at the semantic level.⁶⁴ Only causative predicates allow event-passives. The (simplified) semantic contribution of causative predicates are (138, e_2 =caused event, e_1 =causing event) and the agent relation (139, being a relation between individual x and event e). All passives include (138). Non event-passives (also constructed from causative predicates) also include (139).

Causative predicates are divided into three with respect to construction of event-passives (140-142). Inherently agentive predicates (140) describe events necessarily performed volitionally by animate entities and, therefore, cannot form event-passives. Even if the agent is left unpronounced in (140), it must be identified in context or inferred from world knowledge.⁶⁵ Semantically, the agent relation (139) must be present in the representation of (140), but variable x may be existentially bound. Agentivity-neutral predicates (141), leave open whether agents are involved. (141)'s destruction might be due to wilful individuals or not. Semantically, these predicates may combine with (139) as in (141a), making them equivalent to (140). Otherwise, as (141b), they fall under the final category of non-agentive predicates (142) which admit no volitional agent, and cannot be combined with (139).

In (142/141b), the relationship is solely between caused (main clause) and causing *events* (optionally represented in *by*-phrases, where phrases such as *an explosion* represent event

63 Terminology varies greatly in this area. 'Event-passives' covers a range of non-active (not necessarily 'passive') forms which may be expressed through verbal morphology e.g. Romance SE, or by change of auxiliary e.g. English GET-passives.

64 Not all languages have event passives e.g. Hebrew (Doron 2003).

65 Givón (1990:567-568) for discussion of factors governing identification of agents in such cases.

nominals. In (140/141a), the relation is between caused event and causer, again represented in *by*-phrases. The distinction is obscured in English by use of the same introductory preposition which is vague, covering both eventitive and agentive possibilities (144). As Marantz (1984:129) notes, English *by*-phrases in periphrastic-passives take numerous readings (148), only one of which can be said to reference an implicit agent.

Table 75

Table 7.5

138	$\lambda e_2 \lambda e_1. \text{CAUSE}(e_2)(e_1)$	Causing~Caused event Relation			
139	$\lambda x \lambda e. \text{AGENT}(x)(e)$	Agent Relation			
		Agent	By-phrases	Additional Relation	
140	Today, a bomb was <u>dismantled</u> in Varna	+	(by experts)	$\lambda x \lambda e. \text{AGENT}(x)(e)$	
141	The spacecraft was <u>destroyed</u> yesterday	\pm	a. (by terrorists) b. (by an explosion)		
142	A whale was <u>washed ashore</u> on the east coast	-	(by a freak wave)		
143	A shot killed the criminal Ein Schuss tötete den Verbrecher	Active			[-Agent]
144	The criminal was killed by a shot Verbrecher wurde durch einen Schuss getötet	...from a gun falling to the floor ...fired by intentional Agent			
145	The criminal was killed by unknown persons with a shot Der Verbrecher wurde von Unbekannten durch einen Schuss getötet	Agent and means separated			[+Agent]
146	Unknown persons killed the criminal with a shot Unbekannte töteten den Verbrecher durch einen Schuss	Active			
147		<i>By-Agent</i>	<i>By-Means-Of</i>	<i>With</i>	
	English	by		with	
	German	von	durch		
	Spanish	por		con	
148	Hortense was pushed by Elmer Elmer was seen by everyone who entered The intersection was approached by five cars at once The porcupine crate was received by Elmer's firm The house is surrounded by trees	AGENT EXPERIENCER THEME GOAL LOCATION			

It is often possible to add an agentive *by*-phrase (145) leading to the original *by* being expressed as *by means of* or *with*. These two interpretations map onto different active sentences: (143, [-Agent]) vs. (146, [+Agent]). German *durch* ('through/by/by means of') is also vague, but in this case, true agents are introduced by *von* and the original *durch* remains

constant. Thus, it is necessary to distinguish between the functions of *by*-agent and *by*-means, even if the surface-form is identical (147). This also holds for Romance (§4.4.4).

Given that the semantic representation of event passives is (138) alone, there is no need to assume an implicit argument, since all necessary information is contributed by the predicate itself. Eventive *by*-phrases (142/141b) are fundamentally different to agentive *by*-phrases (140/141a) in that they are not arguments from a semantic point of view. They are simply modifiers of the single event. The fact that they use the same preposition in some contexts in some languages merely obscures this fact. Below, we review the putative evidence for such an implicit argument.

4.4.2 Control

A (c)overt external argument may control into purpose-clauses from periphrastic-passive main clauses (149,151), but not anticausatives (150) or middles (152). This is seen as evidence of a covert intentional animate implicit argument *within* passive, but not middle/anticausative constructions.

Control into purpose-clauses does not necessarily indicate external arguments *within* the host clause. In (153, Williams 1985), PRO may be read as referencing a purposeful controller (evolution/God) not represented in the linguistic structure, often not even in current discourse, but drawn from world knowledge. Similarly in ‘director-contexts’ (154, Fellbaum & Zribi-Hertz 1989) where PRO references the play’s director, finding its referent in that subset of world knowledge pertaining to plays. Nor do purpose-clauses require their referent to be intentional (155). Often PRO references the main clause’s inanimate subject which is clearly

not an implicit agent (155). Williams (1985) and Lasnik (1988) show that PRO need not be controlled by an NP at all, but rather by events. In the acceptable cases of (149-155), PRO can read as referencing the event of the main clause as its subject, unlike (150/152)'s internal events which have no external consequences. Finally, it should be noted that unaccusatives (154) show similar properties. We conclude that PRO's value may be inferred from whatever is referentially available; such examples cannot be used as evidence of implicit arguments even in periphrastic-passives.

Table 76

	Discourse referent	Extern. Ref.	Subj. Ref.	Event Ref.	Main Clause
149	X _i ... The vase was broken [PRO _i to awaken a sleeping child]	✓	✗	✓	Passive
150	*The vase broke [PRO _i to awaken a sleeping child]	✗	✗	✗	Anticausative
151	X _i ... The bureaucrat was bribed [PRO _i to avoid the draft]	✓	✗	✓	Passive
152	*The bureaucrat bribes easily [PRO _i to avoid the draft]	✗	✗	✗	Middle
153	X _i ... Grass is green [PRO _i to promote photosynthesis]	✓	✗	✓	Stative
154	X _i ... The princess dies at the end [PRO _i in order to shock the audience]	✓	✗	✓	Unaccusative
155	Plants _i grow upwards [PRO _i to reach the light]	✗	✓	✓	
156	X _i ... The potatoes are peeled _i [after {PRO _i / _j /our _i } boiling them]	✓	✓	✗	Passive
157	The potatoes _j are peeled [after {PRO _i being/they _j are} boiled]	✗	✓	✗	
158	X _i ... The potatoes peel easily [after {PRO _i / _j /our _i } boiling them]	✓	✓	✗	Middle
159	The potatoes _j peel easily [after {PRO _j being/they _j are} boiled]	✗	✓	✗	
160	X _i ... Babies often roll/turn [after PRO putting them in bed]	✓	✗	✗	Anticausative
161	X _i ... Glasses sometimes break _i [from/after PRO _i polishing them]	✓	✗	✗	
162	X _i ... The boat sank [after PRO _i /our _i putting out to sea]	✓	✗	✗	
163	X _i ... The boat _j sank [after PRO _j putting out to sea]	✓	✗	✗	
164	[IT] La terra si è preparata prima di PRO seminare	~	*Si è preparata la terra prima di PRO seminare		
165	[SP] La tierra se prepera antes de PRO sembrar	~	*Se preparon las tierras antes de PRO sembrar		

Adjunct clauses of passives/middles may license PRO-subjects (Stroik 1992, Reinhart 2000) e.g. (156,158) where *peeler* and *boiler* may be the same person, but are not necessarily so. Such coreference is said to be impossible in anticausative (160) which cannot mean that the children are rolled by those who put them in bed. In (161), however, *polishers* and *breakers* are almost guaranteed to be coreferent. Just as with purpose-clauses, PRO may associate with a discourse relevant party or not (156,158), or with the host clause's syntactic subject

(157,159). Similarly for anticausatives (162,163).⁶⁶ The results of the two sets of examples are clearly inconsistent. If the availability of PRO for purpose-clauses were evidence that passives but not middles/anticausatives possess implicit arguments, then the later examples would seem to prove that middles/anticausatives and unaccusatives also have implicit arguments which would undermine the very concept of unaccusativity. As with purpose-clauses, inference is being confused with hidden local syntax.

Under our definition, non-active clauses do not introduce new agents (covert or otherwise). PRO, therefore, has anaphoric access to agents of prior discourse, or world knowledge should no suitable referent be found (see §5.1.2 for discussion of layered access to referents). Given that the cause(r) will often be present in prior context, its occasional coreference with PRO, is hardly probative of its covert presence *within* the purpose-clause's host.

In passive/middle constructions, the nature of the verb leads us to expect a cause(r)'s existence. Selection of non-active constructions conveys to the hearer that the speaker considers cause unknowable and/or irrelevant to the construal. Nevertheless, the hearer may easily infer them from context or world knowledge (not necessarily matching the view of the speaker). They have no argumental status and are merely attracted to event modifiers.

In fact, a PRO analysis does not appear to be a valid option for Romance. If the main clause is a SE-middle/passive where the subject remains in S_L, that entity *cannot* control into final-clauses (164-165). If this were a case of control of PRO, it would be necessary to conclude that such participants when remaining in S_L were not main clause subjects, which verbal

⁶⁶ Given the disjoint nature of the introducer *after*, event referencing is unavailable.

agreement ensures they are. A more appropriate analysis would be that final-clauses are not active infinitives+PRO, but passive infinitives: ‘The land is prepared before being sown’. If this were the case, non of this putative evidence would be even relevant.

4.4.3 ‘Agentive’ Adverbs

The presence of intentional agents may be implied by other means, such as the prenominal modifier ‘accurate’ in active (166) and passive (167). However, referents so introduced do not have the status of arguments as those introduced by full NPs, and cannot be referenced by pronouns in subsequent sentences (Kamp & Roßdeutscher 1994). Similarly, for ‘agentive’ adverbs. In (168a), the adverb can be said to reference a [–SPEC] agent. This agent cannot be referenced by following [+SPEC] pronouns (*they*), but can be inferred on following [–SPEC] pronouns (*some people*), although the two sets are not necessarily identical. When relevant [+SPEC] agents are already in context (168b), they may be inferred upon the adverb, and referenced as normal in subsequent sentences. (166-168) demonstrate the possibility of referencing agents “at some level of representation” (Hale & Keyser 1986), but that level is not covert local syntax. These are purely contextual inferences.

Table 77			
166	An accurate shot killed the criminal		
167	The criminal was killed by a extremely accurate shot		
168	<table> <tr> <td> John arrived to find a real mess All the windows had been deliberately_i broken ...*They_i enjoyed causing damage ...Some people_j enjoyed causing damage They_i is not necessarily the same as people_j </td><td> The vandals_i had been busy All the windows had been deliberately_i broken ...They_i enjoyed causing damage They_i=vandals_i </td></tr> </table>	John arrived to find a real mess All the windows had been deliberately _i broken ...*They _i enjoyed causing damage ...Some people _j enjoyed causing damage They _i is not necessarily the same as people _j	The vandals _i had been busy All the windows had been deliberately _i broken ...They _i enjoyed causing damage They _i =vandals _i
John arrived to find a real mess All the windows had been deliberately _i broken ...*They _i enjoyed causing damage ...Some people _j enjoyed causing damage They _i is not necessarily the same as people _j	The vandals _i had been busy All the windows had been deliberately _i broken ...They _i enjoyed causing damage They _i =vandals _i		

Periphrastic-passives (169) allow ‘agentive’ adverbs while anticausatives (170) do not. However, such adverbs are *not* incompatible with unaccusative syntax e.g. (173-174, Folli & Harley 2004:47; see also Kallulli 2007), which show unaccusative *cadere* and *rotolare* continuing to exhibit characteristic *essere* selection (typical of unaccusatives) even when the subject acts on purpose. Thus, (170)’s restriction cannot be derived from the unaccusativity of anticausatives. Rather, it is based on the nature of the verb and subject. These verbs are compatible with readings of internal and external causation. In (173-174), the animate subjects can act intentionally even if the event is internal. In (170), inanimate subjects are incapable of intention, leading to the assumption of an external causer which does not match the verb’s internally-caused interpretation. The result is that the sentence is interpreted as passive, which is possible in Romance, since passive and anticausative can take the same SE+verb form, but not in English which requires *was broken* vs. *broke*.

Table 78

169	The vase was broken (on purpose/carelessly)			Passive	English
170	The vase broke (*on purpose/*carelessly)			Anticausative	
171	The vases break easily			Middle	
172	These books read _{i,j} easily [for little children _j]				
173	Gianni {é caduto/*ha caduto} apposta			John has fallen on purpose	Italian
174	Gianni {é rotolato/*ha rotolato} giu apposta			John rolled down on purpose	
175	Los jarrones se rompieron a propósito	}	}	Passive	Spanish
176	Los jarrones se rompieron				
177	Los jarrones se rompieron por sí mismos			Anticausative	
178	Los jarrones se rompieron fácilmente			Middle	

Thus without modification, (176) is read by default as an anticausative, although a passive reading is possible with contextual support. In the presence of *a propósito*, only the passive reading is possible (175). Conversely, the adverbial *por sí mismo* forces an anticausative reading (177). The phrase *a propósito* is a means/manner adverbial which by inference

implies an Actor, but does not necessarily require its presence *within* the clause. Even reference via *by*-clauses does not do this. *Por sí mismo* references the inanimate subject of anticausatives, not external agents (as indicated by agreement). Again, *por sí mismo* is a means/manner adverbial (\approx ‘unaided’). *By*-phrases indicating agents are disallowed (§4.2.6).

English disallows *by*-phrases in middles and anticausatives (Baker *et al.* 1989), but English middles do license *for*-phrases (172) which Stroik (1992) argues denote implicit external arguments. Clearly, the *children* are benefactees of the event. They might also be the *readers*, a fact drawn *not* from the clause, but from prior discourse and/or world knowledge. See Ackema & Schoorlemmer (2005) and references therein for a critical discussion.⁶⁷

Thus, event-passives can take (1) ‘agentive’ adverbs→‘passive’ reading, (2) ‘de-agentive adverbs’→‘anticausative’ reading, and (3) generic manner adverbs→‘middle’ reading. Such implications may clash with the nature of the underlying verb and/or context/participants, creating semantic dissonance, often termed ‘ungrammaticality’, but this is solely based upon the nature of verb/participants and availability of suitable discourse referents. Analyses based on syntactically active implicit arguments, may work in some cases, but purely fortuitously.

Like the reflexive/reciprocals (§4.2.1), a single vague meaning of ‘caused event’ exists for all non-active constructions. The particular reading is derived from context, i.e. combination with other *visible* predicate elements and available inferences. The presence of these adverbials indicates that the speaker does not believe that default interpretation will lead to the listener achieving his/her intended construal. They cannot be used as evidence of covert arguments,

⁶⁷ Note that Spanish *for*- and *by*-phrases use the same *por* preposition.

the presence of which would be shared speaker~hearer knowledge as part of the default reading. The fact that such adverbs can be introduced in order to override the default strongly implies that these putative arguments are not implicit i.e. are not part of that default.

4.4.4 *By*-Phrases

SE-passives with *by*-phrases are found across Romance e.g. Canadian French (Authier & Reed 1996), Italian (Cinque 1988), and Spanish where most “exceptional” cases with *by*-phrases are found in juridical texts (RAE 1973:§41.6).

Table 79

179	El futbolista firmó los contratos	The soccer player signed the contracts
180	Los contratos fueron firmados (por el futbolista)	The contracts were signed (by the soccer player)
181	Los contratos se firmaron (*por el futbolista)	
182	Este cuadro se pintó por {un experto retratista/*Goya}	This painting was painted by {an expert portrait painter/*Goya}
183	Ya habiéndose acordado por el ayuntamiento la inclusión	Its inclusion having been agreed by the town hall
184	Se vigilaba a los prisioneros por los negros	The prisoners were guarded by the Negroes
185	Se ha producido por Nacho Solozábal	It has been produced by N.S.
186	Las pirámides se-construyeron por esclavos	The pyramids were constructed by/with slaves
187	Este país se construyó por mucha gente trabajadora	This country was built by many working people
	*Este edificio se construyó por muchos obreros	This building was built by (the effort of) many workers
	*La basílica se construyó por Miguel Ángel	X was constructed by M.

It is imperative to distinguish periphrastic- from SE-passives. Romance periphrastic-passives, like English, accept a wide range of *by*-phrases (180), but non-active SE-constructions (181) do not admit references to clearly defined agents (see §4.6 for further contrasts). With referential nouns, *by*-phrases are ungrammatical. With [–SPEC] referents, grammaticality often improves (182-183), although judgements are not uniform. Sánchez López (2002:59-61) notes that such forms are indeterminate plurals or abstract/non-specific entities with ‘type’ interpretation. Ungrammatical cases are exactly those where specific agents are present, explaining the variation in acceptability in (187).

Por-adjuncts with SE-passives are rare.⁶⁸ A few analysts accept some examples whilst considering them inelegant (Hernández Alonso 1966:52). Most simply reject them as incorrect; “anomalous/deviant” hyper-corrections by analogy with periphrastic-passives (Luján 1990:97). Arce (1989:199) argues that these *por*-phrases are not agents, but represent means. Similarly, Lenz (1935:96) understands that in (184), the agent is the authority giving orders, whilst *los negros* are the means by which they are executed. CREA provides only one example with an apparent agent (185), but even here, the syntagm may be seen as over-seer of processes executed by others. In (186, De Mello 1978), the slaves are not volitional, but coerced. The causer is the owner who puts them to that task; they are merely instruments.

It is generally accepted that *par*-Agent is ungrammatical in present-day French (Stéfanini 1962; Lagae 2002) but was previously grammatical (Brunot 1965). Heidinger & Schäfer (2008)’s diachronic study of French SE+V+*par* constructions found only 11 examples of potential *par*-Agent in a large corpus covering 1500-1980. If *se*-passives were equivalents of periphrastic-passives, there should be no restrictions on the semantic role of the external argument, but this was the case even during the 1500’s, when *par*-Agent was at its height. All the examples given are amenable to a ‘means’ interpretation. Thus, contrary to common belief, *se*-passive like *se*-middle and *se*-anticausative never license(d) *par*-Agent, but rather *par*-Means which may reference animate but not wilful entities.

Unlike periphrastic-passives, ‘agency’ resides in the deleted causing/facilitating clause. It may be inferred from context, but is *not* a covert part of non-active constructions.

⁶⁸ Reflexive-passives: 0.09% (mostly in Hispano-America). Periphrastic-passives: 30% (Sepúlveda Barrios 1988). Usage has been related to stylistics e.g. legal/administrative language and some periodicals (Contreras 1964:102; Gómez Torrego 1992:28-29).

4.4.5 Other Prepositions

Periphrastic-passives denote a change in viewpoint centred around participants (188). All other participants remain unchanged, because their relationship to the event does not change.⁶⁹

Whilst active~periphrastic-passive alternations operate over participants within an event, eventitive alternations operate over events i.e. between causing/facilitating~resulting events. Events may share participants, but their roles remain separate within each. The intention of the subject of one event is irrelevant to the other event. It is only at the level of the combined predicate that overall intention can be calculated. Since agents can act with/out intention, the same readings are available in composite predicates (189-190).

When the causing/facilitating event is demoted, the function of each participant within its event remains constant, but the relationship between events is determined by linking prepositions (191-193), which select for particular items: *when/after* select events (191), *from/due* to select event nominals (192), both of which may be further elaborated internally. Thus in (192), *applied by John* is an optional adjectival phrase describing the pressure which could be replaced by, for example, *the enormous pressure*. The appearance of a [\pm intent] agent is not evidence for an covert external argument in E₂. The intent of participant within each event remains isolated, and only calculable at the combined predicate level.

By-phrases introducing nominals are unacceptable (193/199) with anticausatives, but marginally acceptable when introducing events (194/200, depending on the particular events being related, cf. 206). Such cases are better with event-introducing prepositions or the full *by*

⁶⁹ Romance does not have 'dative' passivization e.g. Z was given Y (by X).

means of.⁷⁰ *By*-phrases are distinct from the other prepositional introducers, in that they introduce means and are adverbials directly modifying the main verb, *not* separate events (201-206), thus aligning them with *by*-phrases in periphrastic-passives.⁷¹

=Table 80

188	[X gave Y to Z] [Y was given to Z (by X)] [X kicked Y on Z] [Y was kicked (by X) on Z]		Ditransitive Monotransitive	Active Periphrastic-Passive Active Periphrastic-Passive
	External-Causation (crack)			Intent
189	[[E1 John _[-INT] applied pressure to the window] and/so that [E2 the window _[-INT] cracked]]			-
190	[[E1 John _[+INT] applied pressure to the window] and/so that [E2 the window _[-INT] cracked]]			+
191	[[E2 The window cracked] when/after [E1 John _[±INT] applied pressure]]			±
192	[from/due to [E1 the pressure ([applied by John _[±INT])]]			±
193	[*by [E1 John/the wind]]			
194	[%by [E1 John _[±INT] applying pressure]]			±
	Internal-Causation (die)			
195	[[E1 John inflicted a wound on Mary _i] and [E2 She _i died]]			
196	[so that [E2 She _i died]]			
197	[[E2 Mary _i died] when/after [E1 John inflicted a wound on her _i]]			
198	[from/due to [E1 the wound ([inflicted by John])]]			
199	[*by [E1 *John/*the wound]]			
200	[%by [E1 John inflicting a wound on her _i]]			
201	[E2 Mary died due to [an overdose]]			
202	[E2 Mary died at [John's hand]]			
203	[E2 The baby stood [by herself] for the first time]			
204	[E2 The baby stood by [E1 holding her mother's hands]]			
205	[E2 The door opened [by itself]]			
206	[E2 The door opened [by [E1 John pushing very hard]]]			

If availability of wilful agents in SE-passives were evidence of active covert arguments *within* those constructions, then it must be concluded that such arguments are also present in anticausatives (189-194) and unaccusatives (195-200), and even middles. Rather, we argue

70 In fact, English children commonly use *from*- instead of *by*-phrases with passives (Clark & Carpenter 1989) even where *by*-phrases are acceptable, exactly because it identifies a causing event without introducing any notions of volitional agents.

71 In Albanian (Kallulli 1999) and Greek (Lekakou 2005), passives, middles and anticausatives employ the same non-active verbal morphology whilst collapsing the distinction between *by-* and *from-* and many *with*-phrases which are available with all three and active-morphology anticausatives. In all four cases, such phrases can only reference a causer who is *not* capable of wilful agency. Wilful agents are only found with periphrastic-passives.

that when wilful agents do appear, they are introduced by the adverbial/secondary clause itself, in which such agency remains. Unlike periphrastic-passives in both languages, the relationship indicated is between caused and causing events, not caused event and causer. Non-active constructions (Romance $SE_{[-E]}$ -constructions) represent solely caused events.

We conclude that none of SE_{ANT} , SE_{MID} , SE_{PASS} entail an external causing argument at the syntactic level. Rather all of these constructions represent (as we will argue below, case variations of) event-passives.

4.5 Non-Actives as a Mechanism

Although English lacks SE , it does possess a similar semantic arrangement. GET-passives are cross-linguistically common (Siewierska 1984), “normally...without an agent” (Leech & Svartvik 1994:330), placing “the emphasis on the subject rather than the agent, and on what happens to the subject as a result of the event” (Quirk & Crystal 1985:161). Both English and Romance display a full range of causative and ergative get-passives (Huang 1999:45).

Table 81

207	Adjectival-passive	GET-Passive	
Predicate-Adjective	It is big	Causative	Mary got them to fire John
Adjectival-stative	It is broken	Causative $[-R]$ +passive complement	Mary got John (to be) fired
Perfect-resultative	It has been broken	Causative $[_{+R}]$ +passive complement	Mary got herself fired
Passive	It was broken (by someone)	GET-passive	Mary got fired

The adjectival-resultative construction in English is a ‘typical’ passive: agent-less with topicalized patient taking nominative case. Many languages develop constructions derived from periphrastic causative constructions (Givón & Yang 1994). In Romance, since the patient was already the grammatical subject in their respective source constructions (for the

causative constructions, subject of the caused subevent, not matrix verb), they predictably developed into promotional passives where the non-agent topic becomes nominative subject.⁷² This pattern is repeated with reflexive-derived non-active forms in Semitic, Slavic, Modern Greek, etc. (Manney 2000). As with the adjectival-passive, the GET-passive coexists with preceding stages of its evolution (207).

Table 82

208

	+ANIM	-ANIM
	Se curaron los brujos	Se venden bien los apartamentos
Reflexive	The sorcerers cured themselves	
Reciprocal	The sorcerers each other	
Middle	The sorcerers get well	Apartments sell well
Passive	The sorcerers were cured	Apartments are sold well

GET-passives display a non-distinct agent-patient single argument. In Old Spanish, surface-forms took several readings (208). In Modern Spanish, reflexive/reciprocal constructions require personal-*a* (*a los brujos*), whilst such readings are unavailable with inanimate subjects. Subjects *tend* to remain in S_L for passives, but rise to S_H for middles (§4.8). This is typical of grammaticalization processes. Functional re-analysis occurs as a spontaneous activity by individuals during communication, as they extend the use of old constructions and/or words to novel contexts. Once commonly agreed, structural adjustment follows, giving rise to more precise (‘iconic’) coding of newer vs. older functions, as two distinct constructions, allowing them to gradually drift apart following their own developmental paths, although always related by virtue of their common origin.

In what follows, we treat non-active SE constructions as roughly analogous to English *get*_{AUX}. Thus the difference between marked and unmarked intransitive maps onto English the vase

72 See §4.6.6 for development of the non-promotional passive SE_{NAR} into SE_{IMP}.

broke~the vase got broken. As noted for the use of possessive adjectives vs. definite articles (§3.2.1), particular uses will not always map exactly between languages. Most importantly, English *got* obscures two readings, being used for change in status (eventive passive) e.g. ‘He got(=was) killed’, and change in condition (middle) ‘He got(=became) angry’.

4.5.1 Romance Development

In periphrastic causatives, the matrix subject brings about a relationship between undergoer and an event in which the undergoer is the subject. Because the undergoer is an argument of the matrix clause, it appears as a clitic at that verb (209-212). This is often confused with clitic-climbing. In this case, the clitic is already at the matrix verb and does not need to be extracted. The distinction between the different causative constructions is often subtle, and is determined by the nature of the components and perspective intended by the speaker. The difference between DAT and ACC undergoer is obscured in Spanish (209-210) since both cases take personal-*a*, but is clear in French (211-212).

Table 83

209	Ø _i <lo _i > dejé <al nene _i > [PRO _i abrir el paquete] _i	I allowed {him/the child _i } [to open the packet] _i	ACC
210	<le _i > Ø _i hice [PRO _i barrer la casa] _i <a María _i >	I made {her/Maria _i } [clean the house] _i	DAT
211	Je <le _i > ai entendu <Paul _i > [PRO _i clacquer la porte]	I heard {him/Paul _i } [slam the door] _i	ACC
212	Je <lui _i > fais [PRO _i traverser la rue] _i <à Paul _i >	I made {him/Paul _i } [cross the road] _i	DAT

In (213), subject_{NOM} forces undergoer_{ACC} into an action_{DAT/LOC}. In (214) subject_{NOM} forces the action_{ACC} onto the undergoer_{DAT}. When the matrix verb’s arguments are pronominalized, they adjoin to that verb. Note that PRO can be controlled by DAT in (214) even though it usually linearly precedes it, because DAT is a matrix verb argument and therefore syntactically higher. When the matrix subject acts upon himself as the undergoer, the clitic will appear as

SE under the appropriate case at the matrix verb. Depending upon the nature of the verb this can appear “somewhat masochistic” (§4.2.4), and many combinations are avoided.

Table 84

Table 6

	Construction	X's role	Y's role	Y's case	E ₂
213	X coerces Y _i [EVENT to PRO _i ...]	Coercer	Coercee	ACC	[Y is effected]
214	X brings [EVENT PRO _i ...] upon Y _i	Inducer	Affectee	DAT	[Y is affected]
215	X lets [EVENT Y ...]	Facilitator	Undergoer	NOM	[Y changes]

↓

216	S _H []	NOM	OBL	DAT	ACC	V]
217	Y []		(X)		t _i =SE _{PASS}	is effected]
218	Y []		(X)	t _i =SE _{MID}		is affected]
219	Y []	t _i =SE _{ANT}	(X)			changes]

Such periphrastic causative constructions introduce cause(r)s which syntactically and semantically dominate their sub-clause. Introduction of any cause(r) into the sub-clause clashes with that introduced by the matrix verb. Individuals ‘forced’ to act within the sub-clause are, therefore, never wilful agents, but instruments of the matrix cause(r). For ‘causative’ constructions to admit internally-caused sub-events, the matrix subject cannot be a cause(r). In (215), the matrix subject brings about circumstances whereby an event (including its own independent subject) takes place. Thus, X neither effects (coerces) or affects (induces) Y to take any external action; in fact, there is no (in)direct contact between X and Y, the relationship is between the two events. Y is not the matrix verb’s argument, there is no sub-clause PRO, but an independent NOM subject. At no level of representation is X ever a cause(r) acting in relationship to Y, nor Y a causee.⁷³

When the causing/facilitating event is deleted, E₂ retains its meaning. Beyond Y, the only participant potentially relevant to non-active constructions is the original role of X, whose existence and function may be inferred from context. If such an entity is not found, world

⁷³ Events may also occur without an implied causing event, as seen in unmarked anticausatives/middles.

knowledge may provide generic candidates. These are exactly the readings inferred upon OBL in non-active constructions i.e. in addition to its usual experiencer/evaluator roles, OBL may take readings of facilitator, inducer, and coercer (§3.3.6). The intentionality of each role is inferred from context and interlocutor viewpoint i.e. a positive evaluation reads OBL as accidentally bringing about the event, a negative viewpoint sees such uses of OBL as a way of denying responsibility for what was probably an intended, at the very least careless, action. These roles map directly onto OBL's usual reading of positive/negative evaluation (OBL_[+E]) and benefactive/malefactive event affectee (OBL_[-E]), as discussed in §3.3.

When OBL is absent, Ø_{OBL} may be interpreted as a [-SPEC] referent. Since OBL is in a high syntactic position, its referent may be 'picked up' by lower adjuncts/adverbs e.g. so-called agentive adverbs, purpose/adjunct clauses, etc. This explains why the referents of such adjuncts are always [-SPEC] (§4.4.3), and cannot coexist with [+SPEC] OBL. The additional readings are not available with periphrastic-passives as the true cause(r) is syntactically present (even if covert), or with *por sí mismo* 'unaided', the function of which is to deny any external cause (§4.2.6).

Without OBL, sequences for non-active constructions are surface-identical. This is not ambiguity, but vagueness. There exists a single meaning which underlies all: [a COS event occurred]. Often the verb type and/or context ensures the intended reading, but when speakers wish to emphasize a particular property as relevant to their discourse, constructions can be enhanced in various ways, e.g. in most cases, *por sí mismo* picks out SE_{ANT} whilst 'agentive' adverbs pick out SE_{PASS}; whilst relative position of clitics shows that SE+OBL must be

anticausative, OBL+SE must be middle or passive. Like reflexive/reciprocals (§4.2), communications are as vague as suits the speaker's purpose, never ambiguous. SE is interpreted as SE_{ANT}/SE_{MID}/SE_{PASS} as required; a process which can be directed by the speaker.

4.5.2 Non-Actives in Contrast

Removal of the causing/facilitating event (making its cause(r) syntactically inaccessible), leaves only the COS: BECOME (undergoer/property belonging to the undergoer, state). Unlike standard approaches, there is no difficulty in incorporating activity verbs of motion (§4.7.3), as long as they represent internal changes in state e.g. from stationary to in-motion.

Table 85

Source	Non-Active COS	Representation	Case
Mary got [John to be fired]	[John got/was fired]	SE _{PASS} +COS (effectee)	ACC
Mary got [John to become angry]	[John got/became angry]	SE _{MID} +COS (affectee)	DAT
Mary got [the vase to break]	[the vase broke/got broken]	SE _{ANT} +COS (undergoer)	NOM
Mary got [Mary to start moving]	[Mary set off]	SE _{ANT} +COS (undergoer)	NOM

Non-active constructions match the relationships seen for reflexives (§4.2.3). In the ACC version, the undergoer is transformed taking on the verb-defined state as *effectee*. In the DAT version, his/her_{DAT} self_{ACC}, or some relevant state-defined property (ACC) possessed by DAT, undergoes a COS by which DAT is *affected*. Unlike the ACC version, there is no requirement that the process comes to fruition. Thus SE_{MID} (DAT) describes an ongoing COS (*he gets better/becomes fatter*), whilst SE_{PASS} (ACC) describes achieved states (*the book was sold/the treaty has become signed*). The form used is determined by the nature of the verb, the undergoer's, and speaker's intentions, just as the selection of periphrastic causative structures. Some verbs may operate both ways. Thus inherently punctual verbs are restricted to SE_{PASS} unless they can be interpreted as a sequence of such events combining to form a progression,

whilst inherently inchoative verbs are restricted to SE_{MID} unless they can encompass the end result of their scale within their description, often through the addition of adverbs. For the NOM version, the undergoer is the site of an internal change. Some verbs may alternate between SE_{ANT}~SE_{MID}, and SE_{ANT}~SE_{PASS}, depending on context. See (§4.8) for examples.

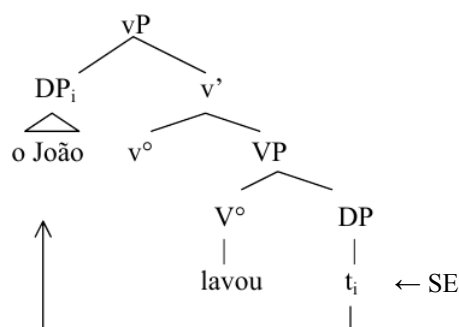
The dative (usually hidden by ACC/DAT syncretism) can be seen in Romanian, and other languages like Icelandic, but there is no one-to-one correspondence between uses of verbs across languages (although they tend to be similar). Moreover, within a language, verbs can appear as both, imitating reflexives where some verbs restrict their patient to ACC, whilst others allow wider range of options, each with its own meaning. Such lexical specifications may also change over time e.g. Old French *aider*+ACC > Modern French *aider*+DAT.

4.5.3 Derivation

In this section, we provide one way in which the derivation of these forms may come about. Other authors propose more sophisticated structures. Hornstein's approach has been selected, purely on the basis that it is the simplest and most diagrammatic. The fundamental point is that the syntactic derivation of non-active surface-forms starts with an arrangement lacking external arguments.

Following Hornstein (2001), a single internal argument DP merges with the transitive verb. Lacking a DP to satisfy the external θ -role, the internal DP moves to [SPEC, ν P], creating a chain with two identical copies and two thematic roles. In Romance, the spell-out for the lower copy (an A-bound trace) is realized as SE, and bears the case of the argument which it replaces: ACC, DAT possessing \emptyset_{ACC} (as for reflexives), or NOM for intransitives (thus

matching the proposed $\text{NOM}_{[+R]}$ in (§4.2.3), and discussed in (§4.7).⁷⁴ When the internal argument is merged, manner/means adverbials have no external argument to reference, leading to the restriction to $[-\text{SPEC}]$ referents (§4.4.3).⁷⁵



The lower copy is $[+R]$ (i.e. coreferent) by definition, but also marked as $[-E]$, since it is not in an ‘external relationship’ to any subject. In addition, t_i retains (or inherits from SPEC, vP) all necessary features i.e. number/gender/person such that spell-out of the pro-DP is a simple ‘look-up’ in the appropriate pronominal paradigm. The difference between case-oriented reflexive $[+R, +E]$ and non-active $[+R, -E]$ is, therefore, based on underlying presence~absence of external arguments, explaining why non-active SE-constructions share so many properties with unmarked anticausatives and middles. In some languages, the spell-out form remains *in situ*, however in most of Romance, clitics raise to positions in IP (or the features are matched there through LDA), where it is case which determines the placement of SE under $\text{NOM}/\text{DAT}/\text{ACC}$ and, therefore, its linear relationship to OBL.

⁷⁴ Many authors, often on very different grounds, have argued that SE must have structural case ($\text{NOM} \sim \text{ACC}$, e.g. Cinque 1988; Dobrovie-Sorin 1998) whilst SE_{IMP} is often referred to as ‘nominative *se*’ (e.g. Oca 1914; Naro 1976; Rizzi 1976). Schäfer (2008:355-368), who employs syntactic structures including voice heads (Harley 1995 and Doron 2003 for similar proposals), shows that Icelandic’s ACC - vs. DAT -marked non-active forms require the existence of at least $\text{voice}_{\text{DAT}}$ and $\text{voice}_{\text{ACC}}$.

⁷⁵ Note that impersonal subjects are merged as $[-\text{SPEC}]$ agents i.e. external arguments. They are not accidents of syntactic derivation, but a positive choice of construal. The fact that in Romance, such agents also surfaces as SE (in most, but not all languages) is a matter of historical syncretism (§2.2.1-2.2.2).

Standard theory relates lexical causative verbs with their anticausative counterparts (Table 86) through semantics such as (222). Whilst theoretical details vary widely, many accounts treat the SE-morpheme in marked anticausatives as reflecting the absence of the external causer argument and the eventuality introducing this external causer argument (the cause predicate and its argument *y*) in the semantic representation of anticausatives (e.g. Grimshaw 1982; Reinhart 2000, 2002; Doron 2003; Reinhart & Siloni 2005; Schäfer 2008; Alexiadou *et al.* 2015). In the approach outlined here, SE does not replace the external argument/causing event, but is a trace of where the sole internal argument was introduced.

Table 86

	Causative	Anticausative
220	Juan rompió el vaso	El vaso se rompió
221	John broke the glass	The glass broke
222	$\lambda x \lambda y [(y) \text{ cause } [\text{become } [(x) \text{ broken}]]]$	$\lambda x [\text{become } [(x) \text{ broken}]] \leftarrow \lambda x <\text{SE} \rightarrow \emptyset> [\text{become } [(x) \text{ broken}]]$

One of the key questions concerning SE is why a historically reflexive marker becomes used in such a wide range of non-reflexive uses. In this model, association at the featural level is very high: (§2.2.1-2.2.2) showed SE spreading along the reflexive paradigm (i.e. overriding person), and across the non-reflexive (i.e. overriding $[\pm R]$) and non-active (i.e. overriding $[\pm E]$) paradigm boundaries. Such featural closeness reflects semantic proximity. §4.6 shows how non-active uses are often reanalyzed as impersonal i.e. lack of an external agent when one is naturally inferred leads to linkage of the form with arbitrary subjects (SE_{IMP}) or situations (SE_{NAR}). For Romance, once Latin's limited medio-passive *-itur* morphology had been replaced at an early date by reflexive-SE with inanimates, then animates (§4.1.2), it was 'free' to spread by analogy in all these directions. Contra Koontz-Garboden (2009, §4.2.6), *all* uses of non-active-SE are *related* to reflexive-SE, rather than just anticausatives. More importantly, they are related, *not* identical.

This raises the question of where the external argument is ‘lost’. It must be different to standard passivization since periphrastic- and reflexive-passives both exist (and have different properties, §4.6.2). The current model assumes that this takes place before the syntactic level.⁷⁶ Thus for periphrastic-passives, the external argument enters the syntax but is removed to an adjunct and the object promoted by syntactic process, as revealed by its accessibility through *by*-phrases and lack of any restrictions upon its nature. Non-active constructions, however, arrive at the syntactic level lacking external arguments, explaining why these are syntactically unavailable, even though world knowledge tells us that they must exist. Reference is only available indirectly as OBL (§3.3.6), or by inference (§4.4). The *raison d’être* of this *lexical* deletion is to show that agents are not *semantically* relevant/appropriate to the construal. In (§4.8), we return to how the limited range of surface patterns which these derivations produce can be interpreted in context.

⁷⁶ For similar approaches see Piñón 2001; Doron 2003; Schäfer 2008; Alexiadou *et al.* 2015. Such analyses do not derive marked anticausatives from their causative variants, rather the lexical derivation of anticausatives from causatives is executed in the Theta System (Reinhart 2000, 2002; Reinhart & Siloni 2005) which is assumed to lack event decomposition; consequently, it does not delete a causative event but only the thematic information about a verb’s external argument/causer (see Horvath & Siloni 2013:218 for details).

4.6 SE_{PASS}~SE_{IMP}

The constructions discussed in this section, loosely referred to as ‘reflexive passives’, have been considered to share impersonality/passivity in contrast to all others, and are often treated as the same item, or at least grouped together in isolation from other uses. We identify three separate (although historically linked) constructions, each with its own clitic/position. Only SE_{PASS} is part of the non-active group being proposed in this chapter.

4.6.1 The Constructions

‘Reflexive passives’ are considered to correspond to periphrastic-passives. Formed from active transitives, primary actants are suppressed (marked by SE), whilst secondary actants rise to syntactic subject position, thereby requiring verbal agreement (223/224). In (225/226), personal-*a* is said to block its rise and thereby concordance, nevertheless SE remains to ‘passivize’ the construction. Since subjects are suppressed, these forms are sometimes considered semantically impersonal. (227/228) are more traditionally impersonal i.e. referencing [–SPEC] subjects (hence, default 3.SG verb agreement) and are available with transitives and intransitives.⁷⁷ The third type, according to our classification (225/226), has been lost in many dialects, but its existence causes difficulties since it can present surface-forms similar to those of the impersonal group when further clitics pronominalize (228).

⁷⁷ Taibo (n.d.)’s statistical survey shows much higher frequency for these forms than their periphrastic-passive and *uno*-impersonal ‘equivalents’.

Table 87

	SE _{PASS}		NP	V
223	Se vende la casa	The house is being sold (by someone...)	SG	SG
224	Se venden las casas	The houses are...	PL	PL
	SE _{NAR}			
225	Se <le> empuja <al niño>	They (people) push him/the boy	SG	SG
226	Se <les> empuja <a los niños>	...them/the boys	SG	PL
	SE _{IMP}			
227	Aquí se vive/come bien	One lives/eats well here...	N/A	SG
228	Non se _{IMP} le _k Ø _j dice cosas _j a mama _k	One does not say such things _j to her _k	SG/PL	SG

Due to overlapping interpretation, some consider (223-228) to be a single (all impersonal or passive) group. Otero (1999:1474-78) considers them to be impersonal. Surface differences derive from mapping to alternative information structures where objects raise to S_H (229), or remain in S_L (230). As Gómez Torrego (1992:29–30) had already pointed out, however, this proposition defines two subjects, a tacit subject external to the verbal syntagm and the explicit one within it, because concordance in (231) proves that the post-verbal syntagm is its subject.

Table 88

229	Ese yacimiento se explotó
230	Se explotó ese yacimiento
231	Se explotaron esos yacimientos
232	{ ^v Se necesitan/*son necesitados} sacerdotes bastante liberales y comprometidos
233	Se hace constar que { ^v se consultó/ ^v fue consultada} a la Excelentísima Corte Suprema

Alcina Franch & Blecua (1975:919) group (223-228) on the basis of shared processes of passivization, whilst Mendikoetxea (1999:170) considers them semantically equivalent; distinctions being merely formal in nature (also Sánchez López 2002:18-35), however, Arce (1989:233) shows that impersonal use of SE with intransitives behaves neither formally nor semantically as a passive. Thus, forcing it into the ‘passive’ group, merely leads to division between two types of impersonals.

Mendikoetxea considers it possible to passivize intransitives and that denial of this possibility is influenced by lack of alternative passive paraphrases. The value of this argument is questionable, since many ‘reflexive passives’ do not allow periphrastic equivalents (232), whilst some impersonals can be paraphrased using passives (233, Taibo n.d.:100-101). Thus, even if suitable paraphrases existed, they would not prove the point which Mendikoetxea desires. Luján (1990:134-148) concludes that shared semantics can only be associations and not structural. It is necessary, therefore, to make divisions on formal grounds.

4.6.2 SE_{PASS}

The traditional term “*pasiva refleja*” (RAE 1973:§3.5.3, our SE_{PASS}) has been criticised as inappropriate (e.g. Seco 1972:119), since this use of SE is not reflexive, and links to periphrastic-passives are indirect. Unlike periphrastic-passives which allow reference to animate subjects (234/235), Modern Spanish SE_{PASS} no longer do so (§4.6.5); (236) is read reflexively: *Pedro se traicionó a sí mismo*. SE_{PASS} constructions allow agreement with other structures functioning as subjects (237). As illustrated by concordance (238/239), these are subject-agreeing passives, and not simply impersonals. (237-239) are unavailable with periphrastic-passives.

Table 89

234	Pedro fue traicionado (por X)
235	Traicionaron a Pedro
236	#Pedro se traicionó
237	Se dice que sin Bizancio el Renacimiento no se comprende (CREA España)
238	Se dice esa verdad
239	Se dicen esos rumores

Position	Determined	Undetermined
Pre-verbal	166	0
Post-verbal	731	397

Taibo (n.d.). Similar results in Barrenechea & Manacorda de Rosetti (1977)

In periphrastic-passives, subjects are usually pre-verbal, but tend towards post-verbal position with SE_{PASS}; necessarily so, if they are undetermined. When determined, position is controlled by discourse factors. Babcock (1970:56) notes that whilst (240) assumes potential viewers, (241) focuses on visibility of the mountains independently of any viewer's presence. In (240), *montañas* comes under the main accent i.e. the information high point is *mountains* not what is done with respect to them. In (241), *ven* takes primary accent so that visibility is the primary information. Compare English *I like Mary~Mary, I like*. The subject's default position is post-verbal, thereby defocusing subjects and presenting propositions as new (Sánchez López 2002:54).

Table 90

240	Se ven las montañas desde aquí	The MOUNTAINS can be seen from here	=You/one can see mountains...
241	Las montañas se ven desde aquí	The mountains can be SEEN from here	=The mountains are visible...

Given that 'reflexive-passives' select different ranges of subjects, convert to instrumental 'through' (not agentive 'by') *por*-adjuncts (§4.4.4), and have the opposite information structure, they cannot be considered semantic equivalents of periphrastic-passives (contra Mendikoetxea 1999:170). In the current proposal, they are eventive passives.

4.6.3 SE_{IMP}

Gili Gaya (1943:§61) maintains that whilst SE_{PASS} is a sign of passivity, SE_{IMP} is an [–DEF, –SPEC] pronoun with significance approximating *alguien*, comparable to French *on*, German *man*, Old Spanish *ome*, Modern Spanish *uno*. Arguments that it functions as subject are also found in RAE (1973:§3.5.6), Oca (1914:573-576), Lenz (1935:§162), and Bull (1965:270). The differences between SE_{IMP} and subject pronouns e.g. *él*, have been amply discussed e.g. Sánchez López (2002:20) and references therein. SE_{IMP} does not allow passive inversion

(242), follows the negative whilst subject pronouns precede (243), cannot be elided for identity (244), nor behaves as a subject in raising (245).

Table 91

242	Se aplaudió a los artistas	*Los artistas fueron aplaudidos por se
243	Uno no debe admirar a los malvados	No se ha de admirar a los malvados
244	Pedro _i sonreía→Ø _i sonreía	Se _i sonreía→#Ø _i sonreía
245	Oigo que se habla	*Oigo hablar a se
246	Se trabajaba en un ambiente tan bueno (Puerto Rico)	
247	Cuando se crece en las calles de una ciudad preñada de violencia, los juegos se vuelven violentos (Spain)	
248	No siempre se es feliz cuando se ama, ¿no es cierto? No siempre se es correspondido (Chile)	
249	De la mujer española se podría estar hablando muchísimo tiempo (Spain)	
250	≈uno podría (inclusive)	
251	≈podrían (exclusive)	

Bello & Cuervo (1960) call them “*cuasi-reflejas irregulares*”. Unlike all other ‘special’ forms, which are restricted in use to specific verb types, SE_{IMP} is found with unergative (246), unaccusative (247), copulars (248), and transitives including those which can/cannot be expressed as periphrastic-passives. It may include or exclude the speaker (249), as seen in paraphrases with *uno*, or 3.PL, and is interpretable as indefinite (*alguien*) or generic (*todo el mundo*). In contrast to ‘passives’, the subject is not suppressed, but prominent. Its key property is simply a non-specific agent (De Miguel 1999; Sánchez López 2002). Whilst it cannot be a subject pronoun like *él*, it may still act a clitic signalling unspecified agents, if we accept a nominative position in structure. (§4.6.8 for comparison between *uno* and SE_{IMP}).

4.6.4 SE_{PASS} ≠ SE_{IMP}

Periphrastic-passives describe states e.g. *the peace was (and is) signed*. The subject (an undergoer) is topicalized by preceding the verb, the state is focused as an attribute applied to that subject, whilst the agent is reduced to an optional *por*-adjunct, retaining its agentive role.

SE_{PASS} constructions are more like middles. As reflected in information structure, SE_{PASS} constructions focus upon *events* modifying post-verbal and, therefore, defocused grammatical subjects (here, *effectee* as opposed to middle *affectee*). Whilst agents are assumed (and often known from context), they are *irrelevant* to the message and unavailable syntactically; optional *por*-adjuncts act as means (§4.4.4). SE_{PASS} highlights the ‘passivity’ of the syntagm *being effected*, rather than attribution of resultant states. SE_{PASS} constructions are *not* ‘passive’ in the same sense as periphrastic-passives.

The ‘impersonality’ of SE_{PASS} constructions is a secondary implication drawn from the subject’s post-verbal position and agent suppression (e.g. Pederson 2005:4-5). The agent’s existence (often identity) is readily available from context; it is simply *not relevant*. Indeed, it conflicts with the message’s purpose of profiling actions *as undergone by* the subject, not actions *taken by* anyone. Mendikoetxea (1999:1643) argues that SE_{PASS} (e.g. *se quemó el bosque para acabar con la plaga de orugas*) reference activities necessitating intentional external agents, differentiating them from inchoatives i.e. SE_{MID} (e.g. *se quemó el bosque*) which are perceived as internal events. Significantly, it is intention/means that is required rather than agent. Such ‘impersonality’ represents lack of interest in, or inappropriateness of mentioning, agents rather than absence. SE_{PASS} is *not* ‘impersonal’ in the same sense as SE_{IMP}.

Conversely, impersonal constructions employ the same information structure as those with explicit subjects. The strong pronoun for this person is Ø in Spanish, but its agreeing clitic appears as SE_{IMP} under NOM. By using SE_{IMP}, the speaker indicates that (s)he cannot specify who the subject is, or uses it to obviate specifying that agent as in normative ‘one does not do

x’=‘you should not do x’. These are not ‘semantically equivalent’ to SE_{PASS}, although some can be paraphrased as such. Although historically related, passives and impersonals (transitive and intransitive) represent two distinct categories.

4.6.5 SE_{NAR}

A key motivation for previous analyses is the inability of SE_{PASS} to combine with animate subjects, leading to employment of the alternative ‘passive’ construction displaying, in our model, SE_{NAR}. Particularly when arguments are expressed as clitics (the functions of which are debated), SE_{PASS}~SE_{NAR} distinctions become easily confused.

In the Old Spanish DOM system, definite human *direct*-objects could be marked with personal-*a* in similar fashion to *indirect*-objects (252),⁷⁸ but need not be (253), whilst *lo(s)/la(s)* were employed as accusative clitics (254). Since then, several regional case-marking patterns have spread obscuring the issue; in particular *leísmo*, where *le(s)* replaces *lo(s)* as direct-object marker (Fernández-Ordóñez 1993, 1999 for details).

Table 92

	Old Spanish	(Examples from Aissen 2003)
252	...rreciba a mios yernos como él pudier major	...receive my sons-in-law as he can best
253	...dexaron mis fijas en el rrobredo	...they left my daughters in the forest
254	Leones lo empuxaron; y el primero...lo comio	Lions pushed him; and the first...ate him
255	Se mataban los cristianos	The Christians were killed/killed themselves/one another
	Modern Spanish	
256	Se mataba a los cristianos	The Christians were killed/ They killed the Christians
257	Se les mataba	
258	Se las mataba (a las niñas)	They killed them (the girls)
259	Se le(la) empuja (a la niña)	They push her (the girl)/ She (the girl) is being pushed
260	Se vende la casa	The house is for sale/one sells the house
261	Se la vende	They sell it

⁷⁸ Romanian (Dobrovie-Sorin 1994a) shows similar properties with preposition *pe* (‘on, upon’ <Latin PER (Holtus *et al.* 1989:104f). For Spanish DOM, see Aissen (2003) and references therein; Laca (2001) for its historical development.

Old Spanish SE_{PASS} was also used with animate subjects (255). Such forms were replaced by SE_{NAR}-constructions (256, systematically from XVI^c) using ‘dative’ clitics for animate participants (257). This construction was gradually replaced after XVII^c with one showing increasing use of accusative clitics. In Ibero-Spanish, this tendency has been particularly strong in FEM.SG/PL *la(s)*, less so in M.PL, and almost absent in M.SG (258-259). Constructions originally based on SE_{PASS} also see an increasing tendency toward substituting NPs with accusative clitics (260-261) in specific dialects (Martín Zorraquino 1979).

The infrequency of animate subjects with SE_{PASS} is traditionally explained as due to ambiguity between such passives and reflexives/reciprocals (RAE 1973:382-383), as found in Old Castilian (Bello & Cuervo 1960:§769; Gómez Torrego 1992:30). This led to the rise of SE_{NAR}-constructions where the preposition marks arguments as ‘objects’. Sánchez López (2002:53-57), however, criticises ambiguity-driven development, as no such restrictions exist in languages such as Italian. Mendikoetxea (1999:1668) links its development with that of personal-*a* which Italian lacks. Nevertheless, verbal restriction to the singular, the argument’s nature/function, or means of commuting forms, are left unexplained. Moreover, particularly in Hispano-America, usage has developed new surface-sequences unavailable in Ibero-Spanish, which are squeezed unconvincingly into one of the existing sets, or where they lack concordance, disregarded as ‘errors’ (§4.6.7).

According to Mendikoetxea (1999:1697-1699), *los* in (262) is ungrammatical in Ibero-Spanish, requiring *les*, which is considered accusative rather than dative (also Fernández Ramírez 1964).⁷⁹ These authors propose paradigmatic explanations where *lo(s)→le(s)* as

⁷⁹ Denied in Bello & Cuervo (1960:§791). Fernández Lagunilla (1975) and Fernández-Ordóñez (1993:78-79) for discussion.

ambiguity avoidance, however, mere ambiguity in such specific contexts is unlikely to have such radical effects. Labov (1994:550) claims that pressure from specific communicative needs is relatively weak, being easily overridden by numerous factors (also Newmeyer 2003); in actual speech, selection of particular variants is rarely the result of intentional individual choice, but forms part of “systemic readjustment”.

Table 93

262	A los herejes se $\frac{\text{los}}{\text{les}}$ quemó	Dialect Spanish	
		Iberian Spanish	
263	se <la(s)> coloca <a la(s) dama(s)>	Laísta Spanish	
264	Se da admiración a Juan	Juan is given admiration	SE _{PASS}
265	Se le [da admiración]	There is a giving of admiration to him	SE _{NAR}
		One gives admiration to him	SE _{IMP}
266	Se le admira	One admires him	SE _{IMP}

According to Alarcos Llorach (1994) *i.a.*, case-marking relates to *ismo*-variations. Studerus (1984) shows that *se+lo(s)* is absent in etymological regions of Spain and Hispano-America, but frequent in Chile and Argentina. However, *le(s)* is also common among non-*leísta* dialects, including non-*laísta* areas of Spain. Alarcos Llorach’s application of *laísmo* to explain impersonals (263) would be “*realmente sorprendente para un hablante que habitualmente no sea laísta*” (Martín Zorraquino 1994:58). Thus, the traditional view that etymological case in the active is directly applicable to SE_{NAR}-constructions is highly problematic, whilst gradual acceptance of accusatives seems to weaken the argument further. It seems unlikely that naturally accusative expressions were commuted to dative in order to avoid ambiguity, only for later generations to reverse the process and reintroduce it.

In fact, diachronic studies (e.g. Bello & Cuervo 1960:§791-792) show that SE_{NAR}-constructions originally controlled ‘dative’ clitics i.e. *se+le(s)* is etymological, not due to *leísmo*. Bello & Cuervo relates SE_{NAR}-constructions directly to ditransitive SE_{IMP} constructions (264-266), but specific developments in peripheral ditransitives seem an unlikely motivator for such large-scale changes. Mendikoetxea (1992:ch.4) suggests that SE is bound to accusative case, perhaps providing pressure for non-reflexive object clitics to take dative case. The argument is weakened by increasing accusative usage whilst SE remains putatively accusative-bound. Furthermore, the clitics would have been inverted (*le_{DAT}+se_{ACC/PASS}*) in this consistently D/A language. Importantly, although Spanish dative-doubling is largely obligatory, it is impossible with these ‘datives’; only when complements have been left dislocated (262), are clitics allowed to fill argument positions. Moreover, the SE_{NAR}-construction is intransitive, “or more accurately, blocks off the possibility of understanding a verb as transitive” (Butt and Benjamin 1994:344), so *le* cannot be an object, direct or indirect. These clitics have dative form but do not function in any way as indirect-objects (even less direct ones).

In our model, the clitic is OBL (which cannot be doubled), whilst SE is the pre-existing SE_{NAR} (e.g. *era+se un rey*, there was a king). This explains clitic order and meaning: *A los herejes se les quemó* ‘there was a killing on (i.e. which affected) the heretics’. It creates a ‘passive’ expression of the killing, treating humans not as objects but event affectees, linking with the deference properties of simultaneously developing personal-*a*, and *-ismos*, whilst explaining the construction’s inherent intransitivity.

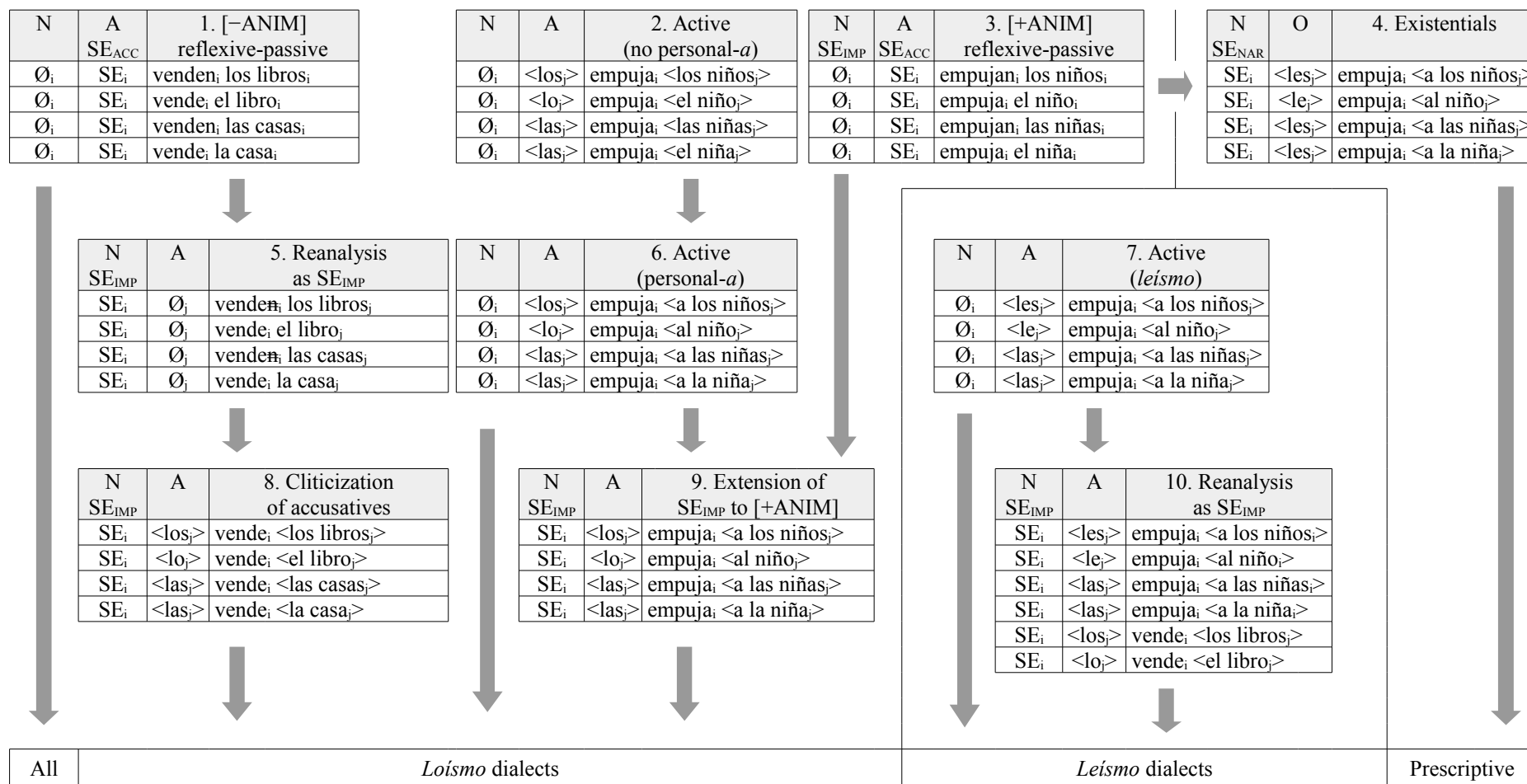
4.6.6 SE_{PASS}>SE_{NAR}/SE_{IMP}

As illustrated in Table 94, the SE_{PASS}>SE_{NAR}/SE_{IMP} development may be seen as successive processes of form-function reanalysis (Labov's "systemic readjustment"):

"Form-function reanalysis is syntagmatic: it arises from the (re)mapping of form-function relations of combinations of syntactic units and semantic components. The process may nevertheless have an apparently paradigmatic result, for example, a change of meaning of a syntactic unit" (Croft 2000:120).

Old Spanish possessed a reflexive passive for [−ANIM] (1) and [+ANIM] (2) subjects, in addition to active (3) and existential (4) constructions. The [−ANIM] passive (1) continues today, but could also be reanalyzed as an impersonal active construction acting upon an [−ANIM] object (5). As indicated by the subscripts and columns, this involves a rearrangement of roles, but the only surface difference is a loss of agreement in the plural, matching that already found in (4). (1) and (5) continue side-by-side as expressive variants. Once established, the new accusatives cliticize as usual (8). The active construction saw the consistent introduction of personal-*a* for [+ANIM] accusatives (6). To this active form with specified subject, it became possible to oppose the indefinite subject established with [−ANIM] objects, i.e. (6)~(9) enter a nominative Ø~SE_{IMP} alternation. (3) can now be reanalyzed as (9) directly, following the same pattern as (1)→(5). The development of (9) can also be seen as paradigm uniformity between (8) and (9), where the accusative paradigms are the same [±ANIM,MASC]=*lo(s)*. For these dialects, the existential form no longer had a function (possibly seen as intrusive *leísmo*), and so falls out of use. By this time, any combination is possible including dative *le(s)* (228, p.169).

Old Spanish



Leísmo dialects saw, in addition to the rise of personal-*a*, a change of ACC.M clitics to *le(s)* (7). Like (6)→(9) above, reanalysis of SE as an impersonal subject produces (10) with its difference from (9) in masculine clitics. For masculine forms, this looks very like (4), and has a similar meaning. Some dialects/speakers adopt (10) which in combination with (8), produces what looks like a single impersonal paradigm with sensitivity to Masculine, Feminine, and Neuter. In this case, (4) becomes marginal. Some other dialects retain (4), making (10) with Feminine referents and (8/9) unacceptable, i.e. (4) is seen as the direct reanalysis of (3).

With the instantiation of [–SPEC] subject SE_{IMP} applicable to any (in)transitive verb, need for SE_{NAR}-constructions fades, as seen in Hispano-America, but in Ibero-Spanish, where SE_{NAR} is retained, SE_{IMP} remains somewhat constrained (e.g. Otero 1968, 1999:1474-1479; Contreras 1964:102-103; Cartagena 1972:117-136). In dialects which lost SE_{NAR}, the new construction developed to take any (in)direct arguments. The case employed is dependent upon the speaker's position along the reanalysis continuum, whilst the forms *le(s)~lo(s)~la(s)* follow his/her dialect rules for each case (*loísmo*, *laísmo*, *leísmo*). Thus, a consistent system of reflexive-passives sometimes ambiguous with reflexives/reciprocals, develops into an increasingly consistent system of impersonals sometimes ambiguous with inanimate passives.

4.6.7 Non-Concordance

Three statistical surveys provide similar results with non-concordance accounting for 10% (Martín Zorraquino 1979), 13% (De Mello 1995), or 9% (Taibo n.d.) of all uses of SE as 'reflexive-passives'. Cases occur with similar frequency in every geographical variety. Given the conclusions of previous sections, classifications for determining these statistics are

probably questionable, and certainly different between each survey. Nevertheless, around 10% across all dialects and registers is too high to be dismissed as ‘errors’.

Taking all forms to be ‘reflexive-passives’, lack of concordance is considered as simply erroneous: “intolerable” (Bello & Cuervo 1960:§792); “*repugna al sentimiento lingüístico del hablante culto*” (Monge 1955:fn.53); becoming “unfortunately” more common each day (Roca Pons 1960:197). For most authors, these form awkward unexplained footnotes. Gómez Torrego and Mendikoetxea consider them ‘deviations’ from passives, but disagree on their characteristics. Gómez Torrego (1992:31-32) considers them (against the trend) to be infrequent in contemporary Castilian but accepted with *determined* nominals e.g. *se alquila estos pisos*, whilst Mendikoetxea (1999:1676) considers them favoured by *undetermined* subjects. RAE (1973) comes closest to the current approach: cases of agreement (*se venden los pisos*) are considered (in the Peninsular, at least) more cultured/literary and read with ‘passive’ significance (\approx *los pisos son vendidos*), while non-concordance (*se vende los pisos*) produces impersonal readings (\approx *alguien vende pisos*).

Table 95

267	PL	PL	Se alquilan cuartos	Rooms are hired	→	(some)one hires out rooms	–ANIM
268	SG	PL	Se alquila cuartos	Rooms are hired	←	(some)one hires rooms	
269	SG	SG	Se alquila uno cuarto	A room is hired	↔	(some)one hires a room	
270	PL	SG	Se alquilan uno cuarto		✗		
271	PL	PL	Se quemeron a los herejes			They went and...	+ANIM
272	SG	PL	Se quemó a los herejes	There was a burning...	=	One burned the heretics	
273	SG	SG	Se quemó al hereje	There was a burning...	=	One burned the heretic	
274	PL	SG	Se quemeron al hereje			They went and...	

The central problem is prior expectation. If all examples are considered passive or active in both form and meaning, then some set of examples will always prove problematic. However, if impersonals and passives are recognized as separate constructions, which as shown above

may generally be paraphrased either way, the problem evaporates. In (267), a passive may imply an impersonal reading; in (268) an impersonal may imply a passive reading; in (269) either is directly available. Only (270) is ‘ungrammatical’, and this is not found with impersonal or passive readings.

40% of all non-concordant cases in Taibo (n.d.) occur in verbal periphrases (275), where lack of concordance is due to the speaker treating infinitive+arguments as the conjugated verb’s complement, instead of constituting a functional unity with the auxiliary. Thus, *puede* agrees with [*poner excepciones*]. There are also sporadic cases of agreement with the ‘wrong’ item. In (276), the verb appears to agree with the direct-object, whilst in (277), Mendikoetxea (1999:59) believes that it agrees with the temporal adverb. In (277), it is clear that ‘Sundays’ do not open, and verbal agreement is with ellipsed subject ‘shops’. (276) may be an example of (cross-linguistically common) agreement-by-sense e.g. English ‘The government is/are deliberating’, where grammatical correctness requires ‘is’ since the government is a singular body, however, ‘are’ is often found agreeing with the plurality of people constituting that body. Once selection between impersonal and passive constructions is taken into account, the number of aberrant cases (only one in CREA (276)) does not warrant the major theoretical debate which it has received.

Table 96

275	Y el propio Gatt ha establecido que se puede poner excepciones (CREA, Chile)
276	Ahí se llevan a los chiquitos que pueden ambular, los llevan y ahí les ponen juegos
277	Se abren domingos (en un local comercial de Valladolid)

4.6.8 Spanish SE_{IMP}

Spanish SE_{IMP} is a [-SPEC] non-reflexive nominative clitic, occupying a row distinct from personal forms, with unspecified number. Contra Rivero (2002) and D’Alessandro (2004) *i.a.*, SE_{IMP} is not equivalent to *uno/la gente*. *Uno* is specific (although [-DEF]), and a full subject pronoun preceding PolP’s negative (279) and positive adverbs (280), whereas SE_{IMP} is in the upper clitic-field following PolP (281/282) (cf. Mendikoetxea 1999 *i.a.*).

Table 97

	S _H	P	N	O	D	A		
278			se _i				lava _i *(a sí mismo) _i	One _i washes _i *(oneself) _i
279	Uno _i	no					lava _i	One _i doesn’t wash _i
280	Uno _i	siempre					habla _i mucho	One _i always talks _i a lot
281		siempre	se _i					
282		no	se _i	me _k	le _j		habla _i así [a la mamá] _j	One _i doesn’t speak _i that way to Mum _j on me _k
283	Ø _i		se _i				duerme _i bien aquí	One _i sleeps well here (SE _{IMP})
284	Uno _i		se _i					One falls asleep well here (SE _{ANT})

Since SE_{IMP} cannot take an object emphatic (278), it is not a VP argument. Since it co-exists with any non-NOM clitic (282/287), and alternates with nominative SE_{ANT} (283~284), it must be NOM. *In vacuo*, the surface forms look like object SE, but may be differentiated by subject specificity. When overt subjects are present (285) including *uno* (286), only specific readings are possible; SE is read as an object reflexive. With no overt subject, the reading derives from contextual specificity of the subject (288-289), defaulting to an impersonal reading (287), where there is no clear subject.

Table 98

	S _H	P	N	O	D	A		
285	Él _i				se _i	las _j	quita _i	[+SPEC,+DEF]He He _i takes _i them _j off (himself)
286	Uno _i							[+SPEC,-DEF]One _i One _i takes _i them _j off (oneself)
287			se _i			las _j	quita _i	[-SPEC,-DEF]One _i One _i removes _i them _j
288			se _i					[-SPEC] book _i People _i don’t read _i much
289		No			se _i		lee _i poco	[+SPEC] book _j People _i don’t read _i it _j much/it _j isn’t read much
290			se _i		*se _i		lava _i	*(a uno mismo) _i *
291	Uno _i				se _i		lava _i	One _i washes _i oneself _i / People _i wash _i

A problem for person-models (noted, but unexplained) is that Spanish SE_{IMP} cannot take reflexive clitics (290). This follows from the case-model: Spanish lacks clitics for unspecified objects.⁸⁰ Since SE_{IMP} is unspecified, the correct output for its reflexive clitic is Ø. Indefinite *uno*, however, being specific, does have a reflexive clitic available (291). By contrast, many Italian varieties possess unspecified object-clitics (e.g. *ci_{IMP}*) and these combinations do appear (§4.6.9). Similarly, SE_{IMP} is mutually exclusive with SE_{ANT} (292-293). This follows from the current model, since both occupy NOM. Moreover, this restriction also affects complements. Whilst [+SPEC, -DEF] *uno* (294) can be doubled by object reflexive complement *sí*, [-SPEC] SE_{IMP} cannot (295).⁸¹ Given the lack of [-SPEC] *sí* as a complement, it is hardly surprising that its clitic form is Ø. There is no such restriction of complements in Italian (296), and hence not in clitic combinations when this complement is pronominalized.

Table 99

292	Mi hermana *(se) desmaya a menudo	My sister often faints
293	*Se desmaya a menudo	Intended: One faints often
294	Uno tiene vergüenza de sí/uno mismo	One has shame of himself/oneself (Otero 2002:172)
295	Se tiene vergüenza de *sí/uno mismo	
296	Quando non si comprende nemmeno se stessi,...	When one does not even understand oneself,...

Surface-oriented approaches employ **se+se*, but cannot explain the phenomena. The above offers a solution based upon observable (and, therefore, learnable) patterns, without *ad hoc* exclusion mechanisms.

4.6.9 Italian SE_{IMP}

In late Latin, SE with ‘middle’ meaning (Brambilla Ageno 1964:201-9) replaced previous ‘passive’ morphology *-itur* (298). Whilst Spanish SE_{IMP} is a recent re-analysis of SE_{NAR} as a

⁸⁰ cf. *Non le_i vió a él_i / *nadie_i*, ‘he did not see me/*anyone’.

⁸¹ Otero (1986:92) argues that Spanish lacks “non-definite objective pronouns” corresponding to English *oneself*, however, since (294) is acceptable, the controlling factor seems to be specificity, rather than definiteness.

NOM clitic, Italian SE_{IMP} has its origin in this earlier process (300, Burzio 1986:43) with uses found in the earliest records (297, Maiden 1995). Classical Latin offers numerous precedents of (298/299), assuming the *-itur~si* correspondence. Such uses were infrequent until the Renaissance. Today, like Hispano-American Spanish, Italian allows all arguments.

Table 100

297	Si può vederli	One can see them	Old Italian
298	Legendo discitur	By reading one learns	Latin
299	Si leggerà volentieri alcuni articoli	One will read eagerly a few articles	Italian
300	<Alcuni articoli> si leggeranno volentieri <alcuni articoli> A few articles will be read eagerly		

Italian SE_{IMP} remains part of VP attaching directly to the verb (Lepschy & Lepschy 1984). Benincà & Tortora (2009) note that SE_{REFL} and SE_{IMP} are not in the same ‘zone’; SE_{IMP} cannot be associated with past-participles (301). The difference between high vs. low SE_{IMP} may be demonstrated by comparing Italian and Spanish under clitic-climbing. SE_{IMP} may appear as the matrix clause subject, but not in any subordinate infinitival clause, where subjects are inherited (302~305). If clitics climb from subordinated infinitives, they appear in their correct positions in the matrix clause, preceding SE_{IMP} under I for Italian (303), and following SE_{IMP} under N for Spanish (306). In personal sentences, object SE_{ACC/DAT} take their normal position (304~307).

Table 101

301

Gli individui [che { ^N si _{IMP} / ^N si _{ACC} } erano presentati presentati- { ^N *si _{IMP} / ^N si _{ACC} }	al direttore] furono...
The individuals [that...	one had introduced (SE _{IMP}) had introduced themselves (SE _{ACC})	...to the director,] were...'

302

[N	O	D	A	I	V ₁	[V ₂	D	A	I]		
Ø _i				<si _i >	può	partire			<*si _i >	One _i can leave	Italian
Ø _i			<lo _j >	si _i	può	dir		<lo _j >		One _i can say it _j	
Ø _i		<se _k	lo _j >	Ø _i	può	mangiar	<se _k	lo _j >		He _i can eat it _j for himself _k	

303

304

305

<se _i >				Ø _i	puede	partir			<*se _i >	One _i can leave	Spanish
se _i			<lo _j >	Ø _i	puede	dir		<lo _j >		One _i can say it _j	
Ø _i		<se _k	lo _j >	Ø _i	puede	comer	<se _k	lo _j >		He _i can eat it _j for himself _k	

306

307

Since SE_{IMP} is available with all verb-types (transitive (311), intransitive (328), copular (314), periphrastic-passives (308)), it cannot be an intransitivizer. Reflexive/non-active uses of SE may be accompanied by subject pronouns; SE_{IMP} may not, but underlying agents are always assumed. In (309) someone is definitely acting, the speaker merely wishes to detach himself from the consequences. Its syntactic equivalence to overt subjects can be seen in control clauses, where subjects do not surface when coreferent with that of the matrix verb (310/311). As a generic subject, SE_{IMP} can be used with adjectives, where it ‘agrees’ with a plural referent (314). Thus SE_{IMP} ≠ *la gente* ≠ *uno* which are singular (312-313).

Table 102

308	Si è giudicati da tutti/dal re ⁸²	One is judged by all/by the king	(periphrastic-passive)
309	Si dice che Giorgio sia stupido	It is said that George is stupid	
310	Lui _i l’ha fatto per Ø _i vedere il quadro	He did it {in order to/that he might} see the painting	
311	Si _i vende _i le scarpe per Ø _i guadagnare denaro		
312	La gente è alta	People are tall	
313	Uno/a è alto/a	One is tall	
314	Si è alti/alte/*alto/*alta	People/One/We are tall	
315	Non <mi/ti/...*ci> si parlò <a mi/ti/...noi> con la dovuta attenzione	One did not speak to {me/you...us} with due attention	

Like Spanish, there is potential for ambiguity (316-318). Addition of object clitics forces impersonal readings (319). Unless left-dislocated, preceding NPs require non-active readings (316/317/320), whilst impersonal or non-active readings are possible when NPs follow. In each case, verbs agree with their subject i.e. following NP (321, passive) or SE_{IMP} (322, active), hence intransitives always take default-person (328). Verbs agree with [–SPEC] SE_{IMP} (i.e. default-person), but adjectives (including compound-tense participles) agree with an understood plural class to which SE_{IMP} refers (323/325).⁸³ Otherwise, participles show subject agreement (326-327). Manzini (1986) proposes that *si*_{IMP} is unspecified for number leading

82 Agent phrase Italian *da tutti* is marginally acceptable, even more so, Spanish *por todo el mundo* (Bolinger 1969). Unrestricted agentive phrases are found in earlier stages of Italian, Spanish, French, Portuguese, and remain available in Romanian (Naro 1968, for the construction’s history).

83 Portuguese SE_{IMP} also only appears with predicates with “group-interpretation subjects” (Naro 1968:12).

tensed verbs to take default 3.SG agreement, and unspecified for number and gender leading to default adjectival agreement which in Italian happens to be masculine plural, while in Spanish it is masculine singular (324).

Table 103

		N	O	D	A	I	Examples from Napoli (1973[1976])			
316	Le porte			si			aprirono		The doors opened	Middle
317					si				The were doors opened	Passive
318						si	aprirono le porte		One opened the doors	Impersonal
319					le	si	aprirono		One opened them	
320	Le aragoste				si		{mangiano/*mangia} d'estate		Lobsters are eaten in summer	
321						si	mangiano le arragoste...			
322						si	manga le arragoste...		One eats lobsters...	
323						si	cantò la canzone <u>tutti</u> insieme		One _{MASC} sang the song all together	
324						si	è facilmente nerviosi		One is easily nervous (M.PL)	
325							è partiti/e presto		One _{MASC/FEM} left quickly	
326							è notato subito le donne		-Agreement => Impersonal	
327	<Le donne>				si _i		sono notate subito <le donne>		+Agreement => Passive	
328						si	va	a teatro	One goes to the theatre	
329	(Io _i)	mi _i	<*ci _j >				pento	<in chiesa _j >	I repent	
330	Gianni _i	si _i					pente		G. repents	
331		ci _i				si			One repents	
332	Di quel peccato _j	te _i	ne _j				penti?		of that sin, are you repenting (of it)?	
333		Ø _i				si	è scritto a qualcuno		One has written...to someone...	
334						si	è scritto e _ito each other... (Reciprocal)	
335				ci _i		si	sveglia di buon'ora...		One wakes up early... (Middle)	
336					mi	si	guidica colpevole		One judges me guilty	
337					lo				...him...	
338					vi	si	guidica colpevoli		...you _{PL} ...	
339					ci				...?us...	
340									We judge ourselves guilty (ci _{1,PL.ACC} +si)	
341									One judges himself guilty (si _{3,SG.ACC} +si)	
342	(Noi,/*voi,)					si	va?			
343	Noi ragazzi,					si	deve...		We boys must...	
344	Noi, non					si	vota per noi stessi		We must not...	
345	Noi,					si	bada alla nostra roba		We pay attention to our belongings	
346	Nous, on						va à Paris?		Shall we go to Paris? (French)	

Its default interpretation as 1.PL is incompatible with 1.PL *ci* when considered as a distinct object referent e.g. (315, Cinque 1988). However, in reflexive and middle contexts (335/336), the reflexive is expressed as *ci*: usually expressed as suppletion $si_{REF}+si_{IMP} \rightarrow ci+si$. The *ci* of

ci+si is clearly not locative/existential (329-331). SE_{ANT} acts as the $NOM_{[+R]}$ of $[-SPEC]$ *si* (331). $SE_{ANT+repentir}$ (332) requires source/cause (*ne*).

For most patterns (336-338) readings are clear, with agreement between adjective and ACC pronoun. (339) is ambiguous. Some find (339) acceptable in the intended reading, but it is generally interpreted as (340/341). This may motivate certain dialect forms, where *noi* (and only *noi*) optionally appears sentence-initially (342), often accompanied by nouns in apposition (343). This phrase is dislocated from the sentence by a pause and is best translated ‘for us (boys), one should...’. When *noi* occurs, SE_{IMP} follows the same patterns and limitations including adjectives agreeing with the abstract subject, hence $\sqrt{noi}/*sé\ stessi$ (344), $\sqrt{nostra}/*propria$ (345). French has parallel forms with *nous/on* (346, Gross 1968).

4.6.10 Other Variations

Rohlf's (1949:234) notes that *ci* is used for SE_{IMP} on the island of Giglio (Tuscany), whilst in many parts of northern Italy, *se+se* is acceptable (347). Others follow Italian's pattern but employ local variants of ci_{LOC} e.g. Vailate (Cremona) *sa+sa*→*ga_{LOC}+sa*, and Neapolitan *se+se*→*(n)ce_{SPUR}+se*.

Table 104

347	...	I	V	...	I	V	One...
Giglio		ci	mangia		si	mangia	...eats
Venetian	se_{DAT}	se	lava	ci_{DAT}	si	lava	...washes
Paduan	se_{DAT}	se	petena	ci_{DAT}	si	pettina	...combs one's hair
Trentino	se_{NOM}	se	'mbarca	ci_{NOM}	si	embarca	...sails (off)
	Dialect Variation			Standard Italian			

Some dialects have developed a Spanish-like high SE_{IMP} e.g. Agliano (348, N. Tuscany, Manzini & Savoia 2005). Many Piedmontese varieties have different forms for reflexives and impersonals (Parry 1998). In Borgomanerese, which is otherwise enclitic (Tortora 2002), impersonal-*sa/as* shares space with SCLs e.g. *a* and may coexist with reflexive-*si* (349-351).

Table 105

	C	N	O	D	A	I			
348		sə		tʃə	nə _i		metta pɔʒə e _i	One puts a little in there	Agliano
349		al					vônga-si _{REF}	He sees himself	Borgomanerese
350		as					môngia bej chilonsé	One eats well here	
351		sa					sta bej chilonsé+si _{REF}	One feels good here	
352				ghe	Ø	se _{IMP}	porta un libro	One brings a book to him	Vicentino
353		se _{IMP}		ghe	Ø				
354				ghe	se		ga presentà	He introduced himself to me	
355		se _{ANT}	ghe						
356					lo	se _{IMP}	vede ingiro	One sees him around	
357		se _{IMP}			lo				
358		se _{NOM}			lo		magna	He eats it (up)	
359				ghe	lo		regalemo	We give it to him	
360		se _{IMP}		ghe	lo		regala	One gives it to him	
361				ghe	lo	se _{IMP}			
362		se _{IMP}		se	Ø		lava le man	One washes one's hands	
363		se _{IMP}		se	lo				
364				*se	lo	se _{IMP}	beve	One drinks it for oneself	
365		sa		ga	Ø		dà al libru	One gives the book to him	Bellinzonese
366				ga	Ø	*sa			
367				ga	la		dò	I give it to him	
368	(*A)			ga	la	sa	dà	One gives it to him	
369	(A)	sa		ga	la				
370	(A)	sa			la		tüt i matin in piazza	One sees her at the square every morning	
371	(*A)				la	sa			
372	(A)	la			Ø		legi, la riviscta	I read it, the magazine ⁸⁴	

Vicentino has developed a high *si_{IMP}* (353/357/362/363), whilst retaining lower *si_{IMP}* (352/356/361/364). Pescarini (2007) notes that these orders are in free variation (independent of socio-linguistic factors) and typical of many Northern Italian dialects. The only oddity in this language (having accepted a case-model), is that **se+lo+se_{IMP}* is unacceptable even though *se_{IMP}+se+lo* is, pointing to a difference between nominative *SE_{IMP}* and that under I. We speculate that the older lower *SE_{IMP}* is 1-person (like Italian), whilst the newer higher *SE_{IMP}* is 3-person (like Spanish) and therefore can display different forms for their reflexives. In the case of lower *SE_{IMP}*, its reflexive would historically be *ghe* (equivalent of Italian *ci*), such that

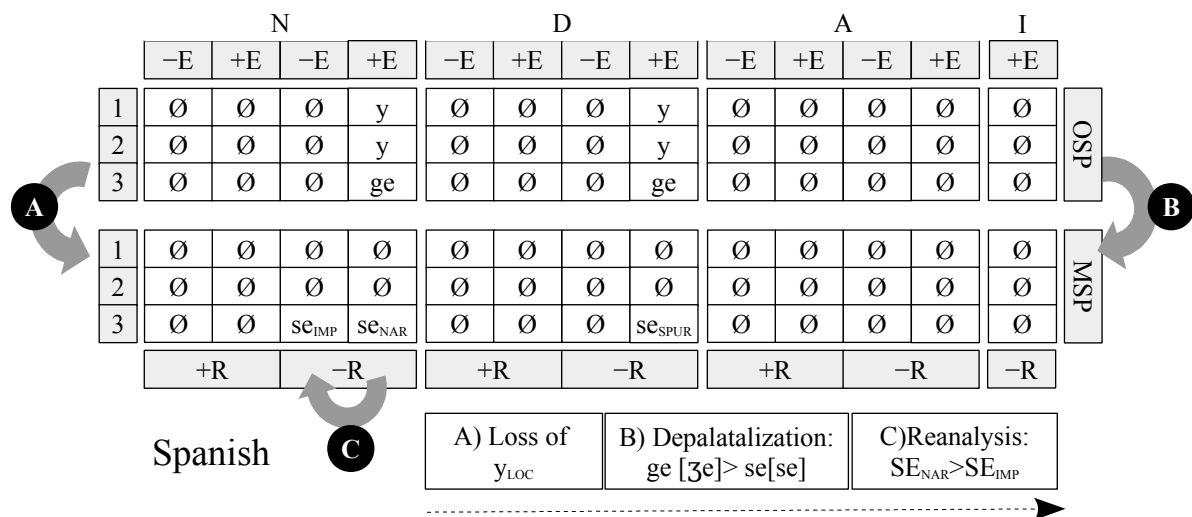
⁸⁴ Cattaneo treats this as an example of *la_{ACC}* moving to the subject position, here it is presented as the standard nominative clitic with ACC-ellipsis.

(364) would give the same output as (361) which is structurally acceptable, although the limited description of the language does not mention whether such an alternative reading is available, or denied due to ambiguity.

Bellinzone (Switzerland, Cattaneo 2009) shows a similar pattern and a restriction $*ga+\emptyset+sa$, even though $ga+la$, $ga+la+sa$, and $sa+ga+la$ are acceptable. The situation is complicated by the fact that Bellinzone displays alternations ACC $la\sim a$, NOM $la\sim\emptyset$, and SCL $A\sim\emptyset$, derived from referent specifications and pragmatics.⁸⁵ It is therefore not clear whence the restriction derives.

4.6.11 Exclusions and Substitutions

In order to cover the range of surface variations, most models require batteries of $*X+Y$ style exclusions with no explanatory power, whilst separating clearly related phenomena. The current model treats these as cases of agreement. The clitic for [DAT/ACC,+R,-SPEC] may be $se/ce/\emptyset$ as determined by the dialect's history. With the additional complication of 3-3-rules in some languages.



⁸⁵ Cattaneo (2009:27-49)'s detailed account of SCL a relates it to (c)overt subjects in the left-periphery.

Early texts show that Spanish had already lost (or had never developed) lower SE_{IMP} and, therefore, had no [−SPEC,+R] counterparts. Loss of *y*_{LOC} and de-palatalization (*ge*>*se*) was followed by SE_{NAR}'s reanalysis as nominative SE_{IMP}, with [−SPEC,+R] forms defined as Ø, like Old Spanish *y/en*. Use with (in)direct objects is a recent development, which may lead to the development of [−SPEC,+R] forms by analogy, but at the moment the DAT/ACC reflexive for SE_{IMP} is Ø, leading to the apparent exclusion.

Old Italian shows lower SE_{IMP} with all combinations of (in)direct-object clitics. Palatalization led to spurious *glie* (§6.2.5), whilst inherited 1/2.PL were replaced by *ci/vi* across (non-)reflexive paradigms of all cases, followed by loss of non-specific and locative *ci~vi* distinctions (§5.2.1). We suggest that this included [−SPEC] clitics. Thus in Italian, the necessary reflexive counterparts of existing SE_{IMP} converted to *ci*, but no high SE_{IMP} developed. It follows that the reflexive of SE_{IMP} is *ci*, hence the putative conversion rule *si+si*→*ci+si*. In dialects where *ci/vi* did not spread it remains *se+se*→*se+se*.

N				D				A				I	
−E	+E	−E	+E	−E	+E	−E	+E	−E	+E	−E	+E	+E	
1	Ø	Ø	Ø	ci	Ø	Ø	Ø	Ø	Ø	Ø	Ø	si	IO
2	Ø	Ø	Ø	vi	Ø	Ø	Ø	Ø	Ø	Ø	Ø	si	
3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	si	
1	ci	ci	Ø	ci	ci	ci	Ø	ci	ci	Ø	Ø	si	IM
2	vi	vi	Ø	vi	vi	vi	Ø	vi	vi	Ø	Ø	si	
3	Ø	Ø	Ø	Ø	Ø	Ø	gli	Ø	Ø	Ø	Ø	si	
1	ci	ci	Ø	ci	ci	ci	Ø	ci	ci	Ø	Ø	si	SI
2	ci	ci	Ø	ci	ci	ci	Ø	ci	ci	Ø	Ø	si	
3	Ø	Ø	Ø	Ø	Ø	Ø	gli	Ø	Ø	Ø	Ø	si	
+R				−R				+R				−R	

Italian

A) Palatalization:
li [li]> gli[ɟi]

B) Replacement:
1/2-person>ci/vi

C) Reducton of
ci~vi distinction

Dialects which developed NOM_[-R] clitics (SCLs) often include [-SPEC] counterparts, whilst early reflexives forms remain available. Vicentino/Bellinzone appear to have retained the early forms whilst creating new ones.

We conclude that both ‘exclusions’ and ‘substitution’ rules of **se+se* type are an artefact of models with too few positions/functions. In our model, they are simply cases of agreement.

4.7 SE_{ANT}~SE_{NOM}

Whilst passives/impersonals are usually separated out (§4.6), remaining uses of SE are generally grouped as showing ‘subject involvement’ at some level. Hernández Alonso (1966:45-50) uses the term *intrinsic*, as opposed to *extrinsic* reflexive uses; another common term is *se-of-interest*, which carries a loose association with ‘ethical datives’. There is, however, little agreement on any further subdivision or terminology.

Fernández Ramírez (1986:§68-69) sees SE as signalling change from the material/concrete to psychological/figurative. Lenz (1935:§158-159) considers intrinsic SE’s separation from true reflexives as “*cuasi insensible*”, merely construing events from an internal perspective. Gutiérrez Ordóñez (1999:1909-1915) considers SE as non-referential, optional, and applicable to any verb type (373-377), because it is independent of verbal valency. Its function is not syntactic, but a marker of focus/emphasis, and unexpectedness. Sánchez López (2002:108-109) considers that it marks an ‘optative’ quality. Lázaro Carreter (1964:389-390) considers SE an affective element which has become attenuated and trivialized by habitual use, whilst Gili Gaya (1964:74) considers them “*excesivamente vulgares*”.

Table 106

373	Pedro se ríe/muere	P. laughed/died
374	Juan se conoce muy bien este país	J. knows this country very well (Imperfective)
375	Nos estamos pasando unas buenas vacaciones	We are having a good holidays
376	María se estuvo callada	M. was (*completely) quiet (=pasar a estar callada)
377	Pedro se supo la lección	P. knows the lesson completely (=pasar a saber la lección)
378	*Pedro se supo que Luis llegaría mañana	P. knows (*completely) that L. will arrive in the morning
379	*Mi hermana se reconoció el error	My sister (*fully) recognised the error
380	*Juan se entregó dos libros a la biblioteca	J. (*absolutely) turned in two books to the library
381	Pedro se comió una cazuela	P. ate up a stew

For transitives, Fernández Ramírez (1986:395) proposes that SE is restricted to transitives with definite direct-objects and “*se acentúa el carácter perfectivo*”, however, neither perfective verb nor definite object are sufficient to make the structure grammatical (379-380), whilst the direct complement need not be definite (381); rather it must be [+SPEC] (Sánchez López 2002:108-9). For Sanz & Laka (2002), direct-objects are incremental themes, whilst SE is a telic marker with properties of delimitation and means, i.e. realization Aktionsart. They criticize (Ordóñez 2002:320) that SE is equally compatible with statives; *saber* (377) is not a predicate of state but realization, because its complement delimits the event, as shown by SE’s incompatibility with *saber* when the complement is unable to delimit *Pedro*’s achievement of knowing (378). Many cases, however, are achievements not realizations, whilst incremental themes which do delimit the predicate, are not always sufficient to license SE (380).

De Miguel & Fernández Lagunilla (2000) argue against SE’s telicity and/or perfectivity. “*Se culminativo*” is an aspectual operator indicating event culmination followed by ingression into a new state, thereby explaining its use with transitives requiring delimitation and intransitives, but also why its unacceptability with perfective (e.g. *llegar, nacer*) or ingressive (e.g. *florecer, hervir*) verbs. Compatibility with stative *saber/estar* (376-377) shows that such verbs can

suppose the existence of previous struggles which have arrived at new states. Affectedness is understood at a pragmatic level, the culmination of a desire (López's optative), with each verb's lexico-semantic properties determining possible readings. The approach does not, however, explain their relationship to anticausatives, middles, etc.

For intransitives, SE is generally treated somewhat superficially. Many note that *ir/morir* maintain different syntax, semantics, and stylistics with their equally intransitive pronominal counterparts e.g. *ir* implies complements of direction, whilst *irse* always requires (c)overt origin (De Molina Redondo 1974:48; Fernández Ramírez 1986:§70; Gómez Torrego 1992:35-36; de Miguel 1999:2986-2987; Alonso & Henríquez Ureña 1971:107). Sánchez López (2002:108-122) considers it to be expletive implying no change in argument structure nor influencing interpretation of participants, but in verbal aspect, equivalent of SE with transitive verbs. This appears to be the consensus of opinion, (Lenz 1935:§160; Alonso & Henríquez Ureña 1971:§129; Manacorda de Rosetti 1961:56; Lázaro Carreter 1964:389; Seco 1972:117; de Miguel & Fernández Lagunilla 2000:13-14; Montes Giraldo 2003:123), but how SE performs these disparate functions, and why only with some verbs, remains unexplained.

Such approaches lead to heterogeneous classes each using SE for apparently different purposes, as already exemplified in §4.3.2. As a result, (Alarcos Llorach 1970:218) opines that it must be purely lexical: “*Su aparición no condiciona en nada la estructura del predicado.*” Lack of syntactic motivation for SE leads to studies concentrating on which verbs can alternate and its semantic effect, however, each author arrives at different sets of meanings, often for identical examples. Proposed categories fail to meet all uses, leading to

inconsistent cases (different for each author) being assigned to the lexicon as irrational ‘pronominal verbs’: *levantar(se)*, *dormir(se)*, *separar(se)* (Gómez Torrego 1992:20–23); *acordar(se)*, *ocupar(se)*, *admirar(se)* (Alarcos Llorach 1970:§5). For Contreras (1964:93-96), SE in *volver(se)* is a lexical diacritic, but indicates distinctions in Aktionsart in *dormir(se)*.

Sánchez López (2002:120) sees SE as expressing change-of-state “intimately tied” to resultant states, echoing Alonso & Henríquez Ureña (1971:106), for whom verbs such as *dormirse/despertarse* signify changes-of-state including a final phase and ingress to a new state. Whilst this makes SE+intransitive similar to de Miguel & Fernández Lagunilla’s *se-culminativo*, there are fundamental differences; intransitive changes-of-state are subject-oriented, whereas *se-cumulativo* is object-oriented. Intransitives are inceptive or durative, transitives are completive. We believe that a key difficulty in understanding intrinsic SE lies in ignoring such differences and its *nominative* status.

In the current model, NOM is an independent position within the syntactic tree (§2.1.1) capable of hosting non-reflexive SCLs, Spanish SE_{IMP}, and even adverbials (§5.4). As a full position, it may also host SE_[+E] (SE_{NOM} as introduced in §4.2.3) and SE_[-E] (SE_{ANT} as introduced in §4.3.4). Below we contrast their functions (as determined by [\pm E]) against each other, and against OBL with which they are often confused. Recognition of these distinctions, not only provides answers to previous issues of classification and functionality, but also allows us to jettison the notion of lexicalized SE (§4.7.5) and special processes related to these items (§4.7.6).

4.7.1 SE ≠ Dative

Traditional grammatical works e.g. RAE (1973:§3.5.4c) treat *se-of-interest* (382) as reflecting an ethical character. Bello & Cuervo (1960:§757-758) call it a “*dativo superfluo*”; but as Fernández Ramírez (1986:395) notes, since some verbs cannot alternate, it cannot be superfluous. Gili Gaya (1943:§58) describes it as “*dativo ético o de interés*”. Alcina Franch & Blecua (1975:914-915) note an intensification of the action. Zagona (2002) considers it a ‘locative’ morpheme signalling co-ubification of predicate arguments, where both suffer transitions coinciding in the event’s final stage. For Gómez Torrego (1992:15-16), *se-of-interest* dispenses functions different to (in)direct objects; it is not ethical but “*una función autónoma*”, which seems self-evident from (402) where it appears alongside direct-object, and ‘ethical’ at the same time. Arce (1989:286) also eschews “*dativo ético*”, calling them “*hipertransitivas*”.

Gutiérrez Ordóñez (1999:1907-15) considers them “*dative reflejo*”, with (383) functionally equivalent to (384).⁸⁶ However, (384)’s most natural reading is malefactive vs. (383)’s agent satisfaction. Dislocated topics highlight the difference: *me*_{OBL} may be doubled, but SE_{NOM} cannot be (cf. Contreras 1964:97; Arce 1989:286). *In vacuo*, NOM looks like benefactive SE_{DAT}, but acts differently. In (385), agent and beneficiary have distinct referents. In (386) they *happen to be* coreferent, thus requiring a reflexive. An agentive reading is also available (387) which can be forced by context (388). When sentences contain both a referentially disjoint PP benefactee and reflexive (389), the latter can only be interpreted as agentive. In (390), *me* highlights subject involvement, whilst *le* is beneficiary. In (391), *me* denotes the internal nature of the process, whilst in (392), *le* introduces an event malefactive (OBL).

⁸⁶ For a similar approach, see D’Introno, González & Rivas (2007).

Appalachian English (Conroy 2007) displays a morphological distinction between agents (396=387) and benefactives (397=386) which also may coexist (398=389). Many authors map SE_{NOM} to high⁸⁷ or low⁸⁸ applicatives, however in each case, SE_{NOM} can be found alongside that applicative, often both simultaneously (393-394). Contra low applicative approaches, ditransitives with SE_{NOM} are plentiful (395, Sanz & Laka 2002, further examples in Gutiérrez Ordóñez (1999:1913).

Perlmutter (1971) considers *te* in (399) to be an *ethical dative*. If *te* were OBL (400), it should read ‘on you’, as (401) reads ‘on me’. In (400), the putative dative is *not* affected by the event, although it might be by consequential actions. Nor can it be an ‘intended affectee’, since this approach cannot then deal with (402) where all positions are filled. Here, *te* is clearly NOM, and yet the ‘intended’ affect remains. What is being signalled in (399) is the agent’s wilfulness (NOM, Arce (1989)’s *hipertransitivas*), not affectedness of third parties (OBL), and it is this that promotes readings with understood consequences. Perlmutter’s example, therefore, must be read as (403) or (404). In fact, the requirement that NOM be reflexive and OBL not be so, is a key means by which these 3-clitic patterns may be successfully interpreted.

87 e.g. Sanz (2000), Sanz & Laka (2002), Borer (2005), Arsenijević (2012), and Boneh & Nash (2011).

88 e.g. MacDonald (2004, 2008), MacDonald & Huidobro (2010), De Cuyper (2006), and Campanini & Schäfer (2011).

Table 107

	Topic/S _H	N	O	D	A		
382		se			Ø _i	bebió una cerveza _i	He drank (up) a beer
383	(*A sí mismo),	se				come toda la comida _i	He at (up) all the food
384	(A mí),		me				
385	Pablo _i			me _k	Ø _j	planchó _i	Paul _i ironed _i some shirts _j for me_k
386	Yo _i			me _i		planché _i	I _i ironed _i some shirts _j for myself_i
387		me _i					I _i ironed _i me_i some shirts _j
388		me _i				... para calmar+me _i	...just to calm down
389		me _i				... por mi mujer	...for my wife
390		me _i	le _j			vestí _i muy bien	I _i dressed _i up for him _j
391				me _i		vestí _i	I _i got (myself) dressed _i
392			le _j	me _i			I _i got dressed _i on him _j
393		te	me	les	Ø _j	cocinaste todo _j	You whipped it all up for them (on me)
394	Juana	se	me	les	Ø _j	bailó un tango _j de miedo	J. danced a beautiful tango for them (on me)
395		se		Ø _i	Ø _j	traía un regalo _j a los nietos _i	He would bring a present to the grandchildren
396	I _i only need to sell me_i a dozen more toothbrushes ⁸⁹						i=AGENTIVE
397	I _i only need to sell myself_i a dozen more toothbrushes						i=BENEFACTIVE
398	He _i went to the store to buy him_i a present for his friend_j						i=AGENT/j=BENEFACTIVE
	S _H	N	O	D	A		
399	Tú	te		me	lo	dijiste	You said it to { me (so you'll have to accept the consequences) me *on you him on me
400			*te	me			
401			me	se			
402		te	me	se			You (went and) said it { to him on me to me (so...) on me (so...)
403		te		me			
404		te	me				

4.7.2 SE_{NOM}

Adding SE to neutral transitive constructions engenders readings of “full exploitation” (Maldonado 2000), where the whole object is physically/metaphorically consumed in a specific time span (408/409); hence (410)’s inadmissibility. For objects to be consumed, they must be totally *effected*, clearly identified, isolatable, and accessible. The object must be bounded and individuated, hence eschewing mass nouns and generics (411b). The contrast parallels English *drink~drink up*, where the particle entails full exploitation.

⁸⁹ cf. French *Je me vends quelques trucs*, ‘I sell me some stuff’ (Boneh & Nash 2011).

Activities (405) may combine with secondary arguments to form *realizations* (406). If fully referential, such arguments delimit activities (acting as measures by which their completion can be recognised) transforming predicates into *achievements*, which *may* be accompanied by SE (407). For Otero (1999:1472) and de Miguel (1999:2995-2997) *i.a.*, SE introduces [+telic] aspect, thereby requiring definite objects, however as illustrated, aspect depends upon the presence/nature of secondary arguments, and it is this *existing difference in aspect* which licenses SE, as shown by its application to *existing* accomplishments. Accomplishments exist for the same verb \pm SE (412a/b). Furthermore, while object restrictions are stringent, aspect is more flexible. Whilst generally perfective (408b/409b), imperfective events are possible (412b). Thus, treating SE as an aspectual operator (e.g. Nishida 1994, Spanish; Roselló 2002, Catalan; and Folli 2005, Italian, *i.a.*), is misleading: SE does not impart aspect, its presence merely indicates when its requirements have been met. Its ‘optionality’ reflects different construals/constructions.

Full exploitation entails subject involvement, extending in some Hispano-American dialects to action verbs (413b). In (414b), *deliberadamente* is acceptable with SE but questionable without it. Equally (415), where the adverbial focuses upon completion. The SE of *aprovecharse* emphasizes subject participation and satisfaction in task completion. Only volition cannot be denied (416). Use of these pronouns is awkward in standard English but is found in some English dialects (Horn 2008). Similar uses are reported in Modern Hebrew (Berman 1981), Arabic (Al-Zahre & Boneh 2010), and Russian (Boneh & Nash 2011).

Table 108

		Second Argument	Predicate Type
405	Pedro (*se) bebe		Activity
406	Pedro (*se) bebe cerveza	–Referential	Realization
407	Pedro (se) bebió una jarra de cerveza	+Referential	Achievement

	(a)	(b)
408	Leyó el periódico con cuidado He read the paper with care	Se leía el periódico de una hora He would read the (whole) paper in one hour
409	Victor sólo comió un poco de carne Victor only ate some/a little meat	Se comió la carne (en tres minutos/#durante una hora) He ate the (whole) meat (in three minutes/over an hour)
410		*Se comió la carne y quedó un poquito Intended: He ate up the meat and some remained
411	La comió despacio He ate it slowly	*Se comió tortillas Intended: He ate up tortillas
412	Bebió un trago a pico de botella He drank a sip from the bottle	Se bebía su tequila antes de comer He would drink (up) his tequila before supper
413		Se bailó una rumba inolvidable She danced an unforgettable rumba (with all her might)
414		Se aprovechó de tu experiencia deliberadamente He took advantage of your experience deliberately
415		Se lo bebió de un trago He drank it in one gulp
416		Me rompí _i algunos coches _j #(sin querer), ¡qué divertido! Vandal: I _i smashed _i me _i some cars _j , #(unintentionally), what fun!
417	*Se miró la tele	Se miró esa película
418	*Se escuchó el murmullo de la brisa	Se escuchó el discurso
419		Se oyó toda una canción de cuna para dormirse
420		Se creyó tus comentarios
421	*Se sabe inglés	Se sabe la lección
422		Se corrió una maratón

423	Me dejé la bolsa en la tienda	I (went and) left the bag at the store	Spain
424	Me olvidé las llaves	I (went and) forgot the keys	Argentina
425	Te perdiste el discurso del director	You (went and) missed the director's speech	Mexico
426	El occiso se entró a su residencia en...	The killer entered his residence in...	
427	Se subió a la silla (de un salto)	He got on the chair (in one jump)	
428	Se subió la montaña	He made it all the way to the top of the mountain	

In others activity verbs, such as directed perception *mirar/escuchar*, second arguments appear to suffer a change quantifiable as consumption (417-418b), but only if the lesser argument makes reference to an entity of delimitation in time, hence (417-418a) are ungrammatical. Equally, state verbs *oír/creer* transform into realizations (419-420). The relationship is

metaphoric of the type “*te trageste todo lo que te dijo*”. Similarly *saber* (421b), where the argument must be completely referential, hence (421a) is ungrammatical. Even simple displacement verbs e.g. *correr/caminar* may express consumption with SE (422), where the distance is seen as being consumed, as seen in metaphors such as “*un auto que devora carreteras*” or “*un bólido que se traga los kilómetros*”.

Acceptance with particular verbs varies across dialects: (423-424) are unacceptable in this construction in Mexico, but are commonplace in Spain/Argentina, whilst only Mexican Spanish accepts (425). Sánchez López (2002:116) denies the possibility of *entrar*+SE, however, it is frequent in various Hispano-American dialects (426, Taibo n.d.:195). Acceptability may even depend on the noun. (427) is acceptable everywhere, but (428) with full exploitation reading only appears in some dialects.

The verb *must* be transitive, either inherently or by virtue of additional elements within the predicate. In our terms, the agent imparts energy into the situation (+E) which returns (+R) as a sense of ‘satisfaction’. In fact, there is a vast literature on the ‘meanings’ imposed by SE_{NOM} which can be contradictory across different contexts. See Armstrong (2013) for a review. In our terms, SE_{NOM}, as discussed for DAT and OBL in §3.5.1, does not carry meaning in itself but is a minimal signal to indicate a significant role for the subject in the construal. Meaning is inferred by the listener from context.

In a case-model, there is no need to ‘calculate’ the features underlying the SE_{NOM} form or move it as an object. SE_{NOM} is simply a nominative reflexive clitic. This approach answers key

questions about the nature of SE_{NOM}, not even addressed by most approaches: e.g. why it has reflexive form rather than another; why it is doubled by nominative emphatics (it has nominative case); why it appears in first position in all clusters (it is merely SE in NOM position); why it is optional (because it is a communicative choice to highlight agentivity in transitive constructions), but enforced with ‘inherent’ reflexives (agentivity is inherent in the root meaning, modulo periphrastic causatives (§4.7.5)).

Since non-reflexive nominative clitics are Ø in most languages, introduction of SE_{IMP} highlights a change from specific to generic/universal. In the case of SE_{NOM}, the effect is to change the focus from the action itself, to the subject carrying out that action. Mentioning the subject in this way invokes a sense of broad ‘subject involvement’, whilst the ‘reflection’ is interpreted from context. From knowledge of a particular agent (likely since this construction is most common in conversation) or people in general (world knowledge), listener’s can reason about the nature and effects of the event as being normal (=>involved/energetic), unusual (=>unexpected), and/or desirable (=>satisfaction). Thus, the Ø~[+R] contrast has the effect of making statements in some way ‘noteworthy’, not in terms of the event itself, but of its contextual evaluation.

4.7.3 SE_{ANT}

SE_{ANT} highlights the pivotal moment of subject-internal change-of-state of (dis)position *levantarse* ‘stand up’, location *subirse* ‘get on’, or translational motion *irse* ‘leave.’ Without SE, these verbs constitute on-going activities. SE_{ANT} is better described as an *inceptive transition into a state* (de Swart 1998), since the focus is upon the *transition* into a new *ongoing* state, rather than the *completion* of the *current* state. Thus, change-of-(dis)position

(*pararse* ‘stand up and *sentarse* ‘sit down’) focus, not on processes of straining muscles, but on the achievement of change-of-state between sitting and standing.

Table 109

	Topic/S _H	N	O		
429		se		murió...después de años de sufrimiento	He _i died _i after years of suffering
430		*se		...en un accidente de coche	...in a car accident
431	Él	*se		murió...suavemente, se quedó dormido...	He died softly, he remained asleep...
432		se		murió...sin que su hijo pudiera hablar con él	He died before his son could talk to him
433	A Juan	se	le	murió su papá	As for Juan his father died on him
434				Un autobús choca en la carretera de Toluca. Mueren 28 personas	A bus crashes on the Toluca highway 28 people die (News report)

Whilst *morir* refers to any death, *morirse* references preparatory phases e.g. an illness (429), incompatible with implications of sudden/accidental death (430, Sanz 2000). *Morir* may represent a natural biological event as an absolute construal without SE (431), or as happening against expectations, directing focus to the pivotal moment marked by SE (432/433). SE’s punctuality is indicated by adverbs (436-439). The central issue is how the event is observed, e.g. (434) where the result, rather than the pivot, is relevant. Such readings are context specific.

Table 110

	Topic/S _H	N	O		
435		se		apareció en el cuarto	He appeared in the room
436	X	se/Ø		despierta diario a las seis	X wakes up everyday at six
437		se/*Ø		despertó abruptamente	X woke up suddenly
438		*se		durmió toda la noche	X slept all through the night
439		se		durmió en clase	X fell asleep in class
440	La lluvia	*se		cae	The rain is falling
441	Adrián	se		cayó	Adrian fell down
442	M	se		cayó de un tercer piso	M. fell (dropped) from the third floor
443		??se		cayó al agua ...con toda elegancia	He _i dived into the water elegantly
444		??se		...vestido	He _i fell into the water dressed
445	La pelota	se		cayó de la mesa inesperadamente	The ball fell off the table...unexpectedly
446		*se		...como era esperado	...as expected
447	La lana	*se		encoge	Wool shrinks
448	El sweater	se	(me)	encogió	The sweater went and shrank (on me)

In (440), rain simply falling cannot take SE, but (441)’s energetic view with Adrian falling suddenly, accidentally and unexpectedly does. *Caer*’s semantics do not allow for agentive expression, so where the diver falls in the water volitionally (443), SE cannot be used. For Maldonado (examples from Maldonado 1988), SE highlights the energy required to effect change. Thus for animate subjects, events are not accidental (444), but necessarily decisive (443). For inanimates, it cannot be normal/expected (446-447), some unspecified force must be exerted (445). Whilst (447) presents the normal state of affairs ([–SE]), (448) has a ‘inceptive’ reading like *morir+se* (429), describing a particular *ongoing-state* coming about. By adding OBL, it may read as a ‘desire’ of the inanimate subject; a form of weak personification. The pattern is quite productive (449-450). Moliner (1984) derives this inference from argument properties. *Caer* occurs in indefinite/non-referential (often generic) contexts (451), and *caerse* in definite/referential contexts (452). Such generic statements are expected, whilst falling events involving definite/referential subjects are one-time occurrences i.e. unexpected, or at least, note-worthy.

Table 111

449	En el parto, la cabeza del bebé fue lo primero que (*se) apareció In the childbirth the head of the baby was the first thing that appeared	+EXPECTED,-VOLITIONAL
450	Juan se (*Ø) apareció en la fiesta sin haber sido invitado Juan showed up at the party without having been invited	-EXPECTED,+VOLITIONAL
451	Caen las hojas en otoño (GEN, NON-REF)	Leaves fall in the fall
452	Se han caído todas las naranjas del árbol (DEF, REF)	All the oranges on the tree fell off

4.7.4 Verbs of Motion

De Molina Redondo (1974:47-56) notes that, for motion verbs, application of SE_{ANT} implies a source (453), otherwise not present (454). In Italian/French, *ne/en*_{ABL} (=Spanish Ø_{ABL}) is required in such circumstances. §5.5.6 provides a detailed investigation of Italian *se+ne* in

relationship to, not only motion verbs, but also stative verbs where *ne_{ABL}* is seen as defining the starting point of the period over which the state holds sway.

Seeing SE as telicity's source, leads to all putatively [+telic] predicates including SE being considered as a class. This results in many (particularly displacement) verbs requiring 'special' treatment, because they focus, not on destination, but on point-of-departure which cannot delimit predicates (e.g. Mendikoetxea 1999; de Miguel 1999:2986; Sánchez López 2002:118). The problem, however, goes deeper. Even when denoting destination, these syntagms do not necessarily delimit the *activity*. Prepositions such as *hasta* 'up to/as far as/for' (455) do not necessarily introduce achievement goals. When such circumstantial phrases represent endpoints, the predicate *may* also be [+telic] but this is context-dependent. As with consumption verbs, it is not SE which introduces telicity, which may not even be present (459-460).

Table 112

453	Ya me Ø _i voy (de aquí _i)	I'm leaving (from here)
454	Pedro irá	Pedro is going
455	Pedro irá hasta la estación	Pedro is going to the station
456	Se fue de la fiesta	He left the party
457	Se fue a Barcelona (para siempre)	He went to B for ever
458	Fue a Barcelona (*para siempre)	
459	Al oírlo se retiró	On hearing it, he backed off
460	Se te ha subido la temperatura ⁹⁰	Your temperature has risen

Since starting-points may co-exist with SE (456), this cannot be SE's contribution either. Indeed, many note that the point of reference is the actor rather than its geographical position. Displacement verbs such as *irse* are not 'special cases' of consumption verbs, but part of the intransitive *morirse* class. Hence in (457), a reading of definitive abandonment is possible

90 Whilst OBL as event affectees is Romance-wide, appearance as clitics is language dependent e.g. Italian *A Giovanni, si {Ø/*gli} rompe il vaso*, 'On G., the vase broke' (§3.3.5).

with SE (unavailable without, 458), not because SE marks a point-of-departure or telicity, but because it highlights a change-of-state in the subject, from being habitually in Barcelona to not ever being there.

Verbs which convert to consumption denote changes-of-state in *external objects*, completion of which defines *achievement*. Anticausatives define *subject-internal* changes-of-state. Whilst both classes highlight subject *involvement*, consumption verbs invoke a pragmatic sense of subject satisfaction (*I ate me a pie*), whereas *morirse/irse* merely indicate that energy has been expended within the subject. In neither case does SE impart any aspectual features. At no time are (in)transitive verbs ‘intransitivized’.

4.7.5 ‘Pronominal Verbs’

Variously termed “*verbos pronominales*” (Bello & Cuervo 1960:§761; Alarcos Llorach 1994:§276), “*verbos de “se” morfológico o estructural*” (Contreras 1964:99-100), “*pronominales puros*” (Sánchez López 2002:96), these verbs do not form a semantically or syntactically consistent class, nor can authors agree on which verbs require lexical storage, since they cannot agree upon the rules to which they are exceptions. From XIII^c-XIX^c, nominative uses of SE became increasingly more frequent, accelerating during XVII^c, in part due to stabilization of personal-*a* (Barry 1987). Bello & Cuervo (1960:§762) postulate an evolution of non→variable→obligatory use of SE, however, its putative ‘obligatory’ nature varies diatopically, diaphasically, diachronically, and even contextually.

Kany (1969) discusses *devolverse* from Chile, Colombia, Costa Rica, Mexico, and Puerto-Rico, which DRAE (2001:810) considers to be exclusive to Hispano-America, developed

from transitive uses by analogy with *ir(se)/volver(se)* (also Gómez Torrego 1994). Similarly limited to Hispano-America (DRAE 2001; Moliner 1967), both forms of *regresar(se)* are frequent. However, while Colombian informants consider that variants may be used freely with little diaphasic or diastratic distinction, SE-variants in Chile and Río de la Plata are less frequent and subject to censure (Taibo n.d.). In some Hispano-American areas (DRAE 2001:1917), *recordarse* is used as a synonym for *despertar*. Moliner (1967:884) considers it to be exclusive to Argentina and Mexico, where it may be used in constructions with direct- (*me recuerdo que una vez...*), or prepositional-object (*me recuerdo de algo*). In fact, CREA provides Peninsular examples (*No me recuerdo cómo se apellidaba*) but none of *recordarse*+prepositional-object, although it was frequent in Classical Spanish.⁹¹ Its presence in the Americas seems, therefore, to be an archaism. DRAE (2001:911) identifies the development of *enfermarse* as an Hispano-American means of emphasizing (de)causative distinctions; *enfermar* ‘make ill’ vs. *enfermarse* ‘become ill’. In Hispano-America, *enfermar* is now considered affected (Taibo n.d.:72), but still occurs, where (contra DRAE 2001) it often lacks a causative reading. Kany (1969) notes its use in rural zones of Spain, again pointing to an archaism, (almost) lost in Peninsular Spanish. Conversely, Lapesa (1981:587) lists numerous verbs with SE in Spain, but not in Hispano-America. Clearly, Bello & Cuervo (1960)’s simple trajectory non→variable→obligatory does not hold.

Latin’s reflexive had not grammaticalized to a middle-marker (Hatcher 1942; Kemmer 1993:161), and ‘reflexive verbs’ are unattested. Middle-marker grammaticalization occurred before the first Old French texts (IX^c-X^c, Stéfanini 1962:583; Kemmer 1993:154), but did not trigger development of ‘pronominal verbs’. Many such verbs arose much later (Hatcher

⁹¹ Kany notes a case in *La Celestina*.

1942:149-202 for numerous examples), even where transitive counterparts existed long before e.g. *se ruiner* ‘to lose all one’s money’ (1559). Similarly, all the verbs identified by Alarcos Llorach (1970) were intransitive in Old Spanish, with SE becoming obligatory only in the XVII^c.⁹² Despite their late appearance, it is the same verbs which end up in this category across Romance. Lexicalization is an unlikely candidate for such parallel development. The reason must derive from each verb’s semantics lending itself to this particular use. Their lexical content has become such that, there are no (or few) situations supporting non-SE use.

Rendir’s original significance was causative, implying that *rendirse* is its inchoative variant. In modern usage, *rendir+se* is obligatory, but is also frequent as *rendir cuentas*. Zero frequency, therefore, does not guarantee that underlying forms do not exist, merely that appropriate contexts are difficult to find.

Alarcos Llorach (1970:216) considers as ‘pronominal verbs’ (e.g. 461-463) only those in which verb+clitic “act as a single element”, functioning as simple verbs e.g. *Juan se queja*=*Juan grita*. Languages with middle systems often have classes of deponent verbs without transitive or intransitive counterparts e.g. Latin *oblivisco-r* ‘forget’ (Kemmer 1993). The group identified by Alarcos can be considered deponents, in that they are inadmissible in any other voice (e.g. **fue arrepentido* ‘was repented’), and highlight active subject participation in emotive actions. Such verbs require SE: one cannot *brag/complain* mildly or without involvement (SE_{NOM}), whilst repentance is an internally-driven COS (SE_{ANT}). Like *enfermar/rendir*, no normal situation allows underlying *arrepentir* etc. to surface. Many verbs have arrived at a stage where non-SE usage is diminishingly small, but rare cases remain.

⁹² Contra Kemmer (1993:160-1), who equates extension of middle-marking with SE’s grammaticalization.

Arrepentir etc. are simply extremes upon a continuum of usage already required for similar verbs, and along which verbs may move over time.

Table 113

461	*(Se) arrepintió de sus tonterías	He regretted his foolish acts	Spanish
462	*(Se) jactó de sus buenos resultados	He bragged of his good results	
463	*(Se) quejó de la política económica	He complained about the economic policy	
464	Ho fatto *pentirsi (a)/pentire Gianni	I made G....repent	Italian
465	Ho fatto *andarsene (a)/andare Gianni	...go away	
466	Ho fatto *uccidersi/uccidere Gianni	#...kill himself/I made someone kill G.	
467	Je fait <i>se</i> _{MID} laver les petits	I make the kids wash/get washed	French
468	Le brouillard fait <i>se</i> _{ANT} humidifier la surface de la terre	The fog makes the surface of the earth humidify	
469	Je fait <i>se</i> _{REFL} laver les petits (l'un l'autre)	I make the kids wash themselves/each other	

There is, in fact, one circumstance in which SE *cannot* appear. In (464), SE is not allowed, but the reading is still available. In (465/466), the reading is not allowed. Control constructions introduce a cause, which is inherited by its sub-clause and cannot be denied (SE_{ANT}) or overridden (SE_{NOM}) by the subordinate verb. Similarly in French (Doron & Rappaport Hovav 2007). SE_{MID} (467) and SE_{ANT} (468) are unavailable when subordinated to *faire*. SE forms can only be read as uniplex events e.g. reflexive/reciprocal (469). This is a Romance-wide phenomenon, whether a verb has accreted SE (and when) in a particular language or not.

We conclude that, like ‘middle’ verbs e.g. *pettinarsi* which are also sometimes treated as lexical units but regularly found without SE, all ‘pronominal verbs’ remain fully compositional. It is merely that the number of situations where they may legitimately be used without SE varies, becoming close to zero for some verbs. The grammatical structures and lexicon inherited from Proto-Romance ensures that even after a millennium of independent development, all these languages will show very similar sets of ‘pronominal’ verbs.

4.7.6 Putative Metathesis

Heap (2005) takes examples such as (470-471) as evidence that **me+se* may trigger clitic metathesis. Both surface forms are determined *a priori* as semantically and underlyingly identical, with surface ‘variation’ requiring explanation.⁹³

Table 114

	N	O	D	A		
470	se _i	me _j			ha escapado _i	It _i got _i clean away on me _j (telic, anticausative)
471	Ø _i		se _i			It _i became _i free on me _j (atelic, middle)
472	{‘Se mos/’mos se} escapa				We’re losing it	Judeo-Spanish
473	El livro podia kayer- {‘se-mos/*mo-se}				The book could fall	
474	{‘Me s’/’mos s’} escapa				I’m losing it	Baix-Ebre Catalan
475	No podia escapar- {‘se’m/*me-se}				This couldn’t get lost	
476	{‘Se me/’me se} escapa				I’m losing it	Murcian Spanish
477	Puede escapar- {‘se-me/*me-se}				I could lose it	
478	{‘Se le/*le se} escapó				It escaped him	

Verbs such as *escaparse*, however, are ‘degree achievement’ verbs (Hay *et al.* 1999), interpretable as telic (470) or atelic (471), leading us to expect two constructions containing SE_{ANT} or SE_{DAT}, as illustrated. (470) focuses upon the pivotal point defining the end of the struggle and movement into a new state of loss. (471) highlights the ongoing struggle itself. *Haber*+PP places both events in the past, but defines neither as perfective; completive~durative are defined by SE_{ANT}~SE_{MID}, which is made clear by OBL.

Ordóñez (2002) reports cases in several varieties (472-478).⁹⁴ *Me*’s OBL status is confirmed by its unavailability following infinitives, and unavailability of non-existent 3.OBL_[+E] (478,

93 This alternation has been evidenced since XIII^c (Lapesa 1980: 472). Heap (2003)’s statistical survey of COSER and ALPI show consistent availability of *se+me/te* and *me/te+se* but at a much lower frequency. This is to be expected since use of OBL with transitives are designed to add immediacy to the statement invoking interlocutor reaction. Such usages are less likely to be documented. In fact, counts only go above twenty per century in the last period where the ALPI project set out specifically to record spoken usage. According to Heap (2003), there are definite register and dialect preferences for some forms.

94 Similarly, Dominican Republic Spanish (Rivera-Castillo 1997).

§3.3.5). Such ‘alternatives’ are semantically distinct construals presented in underlyingly different syntax, as revealed by OBL when present. Far from requiring complex rules, such forms are evidence for a *simpler* underlying structure leading to iconic representation.

4.8 Composition and Interpretation

The previous sections have presented an array of constructions, all of which surface as SE+verb, with multiple potential readings. Throughout, there have been three key indicators as to the most appropriate reading: information structure which indicates the level of subject agency/dynamism; knowledge of subject capabilities (as discussed in §4.3.3, and largely reflected in its animacy); and the nature of the verbal root itself.

Within the non-active group, the central participant is an agent in terms of its “teleological capabilities”, but not dynamic. With SE_{MID}, subjects tend to rise to S_H, indicating their involvement in the development of what is an inherent property; with SE_{PASS}, subjects tend to stay in S_L, underlining their lack of dynamism as an (often non-inherent) property is applied to them. SE_{ANT}, tends to prefer S_L. Either tendency can be overridden for pragmatic purposes. Subjects merged at S_L only raise to S_H if they are agentive. Thus pre-verbal position strongly implies middle i.e. topic (=subject)+comment (=attribution of properties) or reflexive readings. Remaining in S_L, allows the same readings but it is more likely interpreted as passive (Mendikoetxea 1998, 1999:1657; Sánchez López 2002:66; Felíu Arquiola 2008). Pederson (2005) notes that semantic impact of position is highly dependent on verbal lexical specifications. The effect is substantial with *abrir/cerrar*, but minimal with *construir/vender*.

With [–human] subjects (479-480), pre-verbal position defaults to middle readings; post-verbal position to passive readings, although either reading is possible in context. A reflexive reading is not possible, since these subjects do not have ‘mental state’. With animates, however, the passive reading is avoided in Spanish, and since they do have ‘mental state’, a reflexive reading is possible, with information structure determining the default reading out-of-context (481-482). Otero (1999:1471) notes that higher animals seen as possessing ‘mental state’ are treated as [+human] and volitional, thus *el gorila se mató* would be treated as (481-482), rather than (479-480).

Table 115

		Default	Possible	Unavailable	
479	Una mosca se mató	Middle	Passive	*Reflexive	A fly got/was killed
480	Se mató una mosca	Passive	Middle	*Reflexive	A fly was/got killed
481	Luis se golpeó	Reflexive	Middle	*Passive	Luis hit himself / Luis got hit
482	Se golpeó Luis	Middle	Reflexive	*Passive	Luis got hit / Luis himself
483	El jarrón se rompió	Passive	Anticausative	*Reflexive	The jar was broken/broke
484	Se rompió el jarrón	Anticausative	Passive	*Reflexive	The jar broke/was broken

		Context	Reading	
485	Pedro se controló	con los años	Inchoative	P. gained self control over the years
486	Pedro se controló	para no asustar a los niños	Reflexive	P. controlled himself so as not to...
487	Se cansa	en la tarde	Inchoative	She gets tired in the afternoon
488	Se cansa	a propósito para dormir mejor	Reflexive	She tires herself purposely to sleep better

Whether a ‘middle’ or ‘anticausative’ reading is available depends on root semantics. With animate subjects, context often determines the reading. Some verbs e.g. *controlarse* ‘gain control’ (485-486) and *cansarse* ‘become tired’ (487-488) generally only allow middle interpretations, but may rarely take (pseudo-)reflexive readings in sufficiently strong contexts.

Some roots are inherently punctual and hence restricted to anticausative readings (e.g. ‘break’), others describe processes and are therefore restricted to middle readings (e.g.

‘anger’). Others can vary between the two (e.g. the ‘degree achievement’ verb *escaparse* discussed in §4.7.6). By virtue of this information, the range of possible readings is limited, indeed often singular in a given context.

4.8.1 Conclusions for SE

A verb’s lexical specification determines how many arguments *must* be filled. Additional participants may be added giving the impression of increasing its valency e.g. monotransitives may receive an additional DAT as possessor of ACC, intransitive activities may take adverbials of measure, ‘pseudo-transitives’ e.g. *run a race*.

Without SE, predicates are neutral with respect in their ‘perspective’ and each participant may be topicalized/focused in various ways. Introduction of SE changes the predicate to one which is viewed from the subject’s perspective. When events are seen as leaving the subject and entering the outside world (either underlyingly transitive, or pseudo-transitive), they may ‘reflect back’ onto the secondary role played by the subject, in which case the predicate is defined as external [+E]. Alternatively, the predicate may be defined as internal [–E] by addition of the other class of reflexive pronouns,⁹⁵ where the event takes place only from the subject perspective and other arguments become irrelevant e.g. we acknowledge agents in passives/middles, they are merely irrelevant, indeed inclusion would clash with SE_[–E]. It is, therefore, possible to ‘internalize’ both transitives and intransitives, without changing their transitive status; middles/passives are not intransitivized transitives, and no complex propositions are required in order to intransitivize intransitives. Nominative SE (SE_{NOM}~SE_{ANT}) can apply to almost any verb, precisely because every verb has a subject.

95 These two sets often have different forms in other languages (§4.3).

Table 116

	–E	+E	–E	+E
NOM	SE _{ANT}	SE _{NOM}	SE _{IMP}	SE _{NAR}
DAT	SE _{MID}	SE _{DAT}	Ø	SE _{SPUR}
ACC	SE _{PASS}	SE _{ACC}	Ø	Ø
	+SPEC		–SPEC	

All uses of SE are compositional, there are no pronominal verbs. Unlike previous classifications, the current model is clear cut whilst reflecting the gradient nature of usage. For the [+E] attribution of SE_{SPUR}, see (§6.2), and for the relationship between SE_{NAR} (<ge) and existential locatives, see (§5.4).

4.8.2 Adequacy of Form(s)

If all these functions took different forms, analysts would have no difficulty in separating them out. However, this would be to ignore the reality of human communication and the history of these particular languages. There is no source for a differentiation between SE_{NOM}~SE_{ACC}; even the Latin distinction between SE_{DAT}~SE_{ACC} has been lost due to phonological pressures (§2.2.2) in all languages except Romanian. It is part of the efficiency of language to transmit the minimal amount of data required for communication, based upon expectation of default interpretation by the listener. If these elements are capable of proper interpretation without the burden of extra forms (as they are), then it would be inefficient to maintain them. Indeed, some languages no longer entirely do so (§2.2.1).

Reflexive and non-active constructions are often vague. The listener is expected to interpret the signals in light of world knowledge, knowledge shared/developed between interlocutors, and the position of the message within discourse. Such interpretations depend for efficiency on default readings. Usually, speakers leave listeners to make the obvious choice of

interpretation. When necessary, speakers guide such interpretation by enhancing the message. This is true efficiency rather than the *a priori* reduction of options discussed in §2.2, and reflects real language use rather than idealised and mechanical theories.

Moreover, these constructions do have different forms when required. The adjuncts added in order to differentiate the constructions do so by indicating differences in case, both positively in terms of form e.g. *lui même~à lui même*, and negatively by denying SE a particular case e.g. the presence of accusatives ensures that the reading cannot be one of SE_{ACC}/SE_{PASS}/SE_{ANT} etc. Because the number of options is fixed, very limited amounts of additional information are required in order to guarantee exact communication. Conversely, in order to support such efficiency, there must exist a set of distinguishable patterns onto which communications may be mapped. In a case-model, this is provided by the system of four case positions which not only imposes interpretative restrictions (contra García, §1.4.2) e.g. OBL differentiates between SE_{NOM}+OBL and OBL+SE_{DAT/ACC} (§4.7.6), but also results in clitic sequences being iconic representations of the construal (§2.1.1).

In vacuo, interpretation of SE is intractable. In context, minimal signals indicating who is related to what, allow ‘meaning’ to be inferred. This is only possible if the parser is aware of multiple targets for the same surface form e.g. OBL can only differentiate middle~anticausative if the underlying model has three potential targets: SE_{ACC}/SE_{DAT}/SE_{NOM}. If all SEs are the same, then all surface-identical forms are underlyingly identical, and such alternations are random, rather than informative.

4.8.3 Adequacy of Model

Contra person-models, where clitics appear in different positions depending upon their neighbour, it is better to have a fixed number of positions sometimes filled with Ø. When we do this, impossible combinations become readily interpretable as natural extensions of those already understood, complexities such as non-active constructions and awkward details such as SE_{IMP} sometimes rejecting object reflexives emerge naturally without the need for any clitic specific mechanisms.

We also *need* this many categories. Without them, we mix up two types of dative, three types of reflexive, and three types of non-active construction, resorting to *ad hoc* rules based upon semi-equivalence of meaning to cover the discrepancies. Without them, it would be impossible to express the range of construals available through such a small number of forms and without this many ‘targets’, a parser could not reconstruct the underlying form from surface-identical forms. Contra García (§1.4.2), we consider structure to be the key element in language which makes interpretation possible.

Once these categories are accepted, the level of ambiguity even in Spanish, with its ubiquitous SE, is unproblematic; the different underlying structures can be readily re-constructed by the listener from identical surface forms *in context*. As discussed in §3.5, this is possible precisely because of the minimality of the signals given and application of a shared inference engine.

4.9 Conclusions

This chapter has identified the *range* of, and *need* for, numerous functions, often confused by virtue of identical form. We distinguish case distinctions for reflexivity (including

nominative), non-active constructions (including not only passive vs. anticausative, but also middle as a separate item), and distinct impersonal constructions. This variety is expressed in terms of the same concepts of ‘case’ and $[\pm E]$, as used for non-reflexives in the previous chapter and non-personal clitics discussed in the following chapter.

5 NON-PERSONAL CLITICS

This chapter considers non-personal (sometimes called ‘adverbial’) clitics which often require more ‘interpretation’ than direct and physical referents. We illustrate that each adverbial clitic has more functions and can appear in more ‘cases’ (and, therefore, in more positions) than is usually understood i.e. can express a wider range of concepts than ‘simple’ clitics which reference objects, with wider reference than physical places. This leads to sequences of clitic+verb taking ‘idiomatic’ readings and discussions of lexicalization. We argue that all such ‘special meanings’ can be identified from, and composed within, syntax. There is no need to treat any such usages as having been removed from language as “unanalysable chunks” (Chapter 1) and, therefore, no need for lexical storage. Rather, we argue that the model predicts, and our analysis supports, a purely compositional approach.

5.1.1 Against Lexicalization

One approach to clitic ‘idiosyncrasies’ is to see development from WPs to clitics as including fossilization of certain combinations, involving “the grammaticalization of the clitic pronoun into an obligatory morpheme, which no longer functions simply as pronominal element... [and]...lexicalization...introduction into the lexicon of the verb+clitic (+adverb/nominal) sequence as an independent item” (Russi 2008:112-3). There is, however, no agreement concerning which combinations require lexical listing (1, from **Russi 2008**, **De Mauro 1999-2000**; **Kinder & Savini 2004**), or explanation of why similar cases remain compositional.

Table 117

1

	R	D	K
averne abbastanza	✓	✗	✓
non poterne più	✓	✓	✗
intendersene	✓	✗	✓
volerne	✓	✓	✓
fregarsene	✓	✓	✓

	R	D	K
infischinarsene	✓	✓	✓
fottersene	✓	✓	✓
sbattersene	✓	✗	✓
impotarsene	✗	✓	✗
andarsene	✓	✓	✓

	R	D	K
venirsene	✓	✓	✓
partirsene	✓	✗	✗
(re)starsene	✓	✓	✗
uscirsene	✓	✓	✗
tornarsene	✗	✓	✗

Such lexicalist approaches⁹⁶ presuppose clear classification of functions available to each clitic, allowing identification of non-adhering cases. But, from our perspective, it is precisely this understanding which is absent. Russi follows Sala-Gallini (1996:87) regarding *ne* as a strictly grammatical element signalling *accusativo genitivale*, as evident from its ‘obligatory’ presence with certain verbs which retain full complements which *ne* is ‘expected’ to substitute (Russi 2008:113). The clitics in question, however, are partitives under ACC, whilst the simultaneous *di*-phrase pronominalizes under DAT; there is no doubling. Calling this *ne* ‘obligatory’ is simply to state that transitive verbs must realize their direct-object. Indeed, Russi (2008:113) notes that “it would be more accurate to attribute this lexicalized *ne* the function of indicating that the object of the verb need not be overtly expressed. In other words, we are dealing with the phenomenon of null-object instantiation”, which is effectively to recognise *ne* as ACC. We argue that recognising each clitic’s multiple functions makes lexicalization unnecessary.

Although this chapter deals with syntax across Romance, it focuses on Italian ‘idioms’ illustrating clitic functions, individually (V+*la*, V+*ne*, V+*ci*) and in combination (V+*cela*, V+*sene*, V+*sela*) demonstrating that all cases are compositional. Many examples are taken from Russi’s work (representing one of the few in-depth synchronic and diachronic studies of such clitics in any Romance language⁹⁷), not only because it represents a comprehensive resource, but also to highlight that it is not different data which leads to our different analysis, but an insight into the multiple range of meanings that each clitic may carry as a result of their fixed properties in relationship to case (as reflected in position).

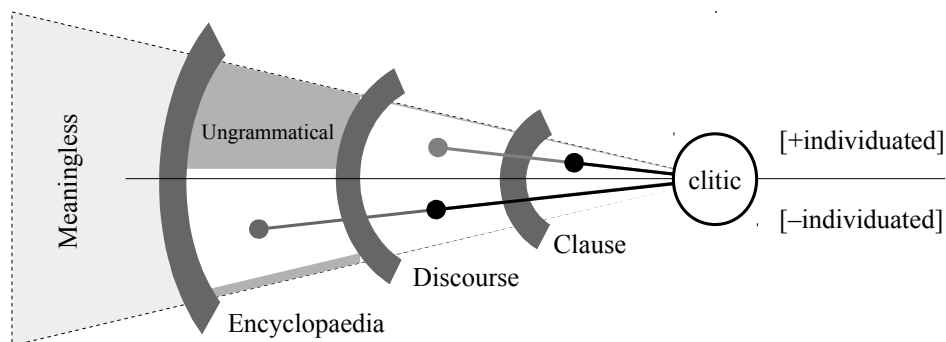
96 See also Abeillé *et al.* (1998) for French.

97 Also Espinal (2009) for extensive Catalan examples.

5.1.2 Interpretation

The second issue is how clitics gain meaning. Many (sometimes termed ‘idiomatic-clitics’, I-clitics) appear to lack referential interpretation (i.e. have no syntactic antecedent, nor correspond to individual entities/locations) and are often presented as cases of lexicalization.

Clitics must combine with information available from discourse to identify conceptual antecedents. Delfitto (2002) represents all Romance clitic constructions as (hidden) left-dislocation constructions; the clitic’s binding-theoretic contribution is a formal object encoding $\lambda x[...x...]$, where λ -abstraction must combine with a (hidden) topic which “counts as the argument of a λ abstract” (Delfitto 2002:52). Hence, Italian *questo libro, l’ho letto*, becomes ‘ $[\lambda x \text{ (I have read } x)] \text{ (this book)}$ ’. I-clitics “give rise to unsaturated λ abstracts” (Delfitto 2002:49), with λ -abstracted variables encoded over a range of non-referential topics (or right-dislocated constituents) denoting an ontology of abstract objects: propositions, properties, generic situations, spatio-temporal locations, or indeterminate objects, depending on the clitic’s properties and the content of the most accessible topic.



Clitics signal that antecedents are highly accessible, even if covert. Processing proceeds outwards. Referents are queried amongst the closest (i.e. clausal) and individuated objects, then wider discourse, and finally encyclopaedic knowledge, guided by the clitic’s ϕ -features.

[+individuated] clitics must be matched within the set of [+individuated] antecedents; failure to do so is ‘ungrammatical’. If clitics can be read as [±individuated], the inner→outer sequence determines that if an [+individuated] match is found, the combination is interpreted directly; hence ‘idiomatic’ meanings become inaccessible in the presence of clear antecedents (3). If no such match is found, an appropriate referent is queried first from within the wider discourse and then encyclopaedic knowledge; failure at this level remains grammatical but ‘meaningless’. Whilst clitic properties remain constant, the most accessible topic changes with discourse, hence interpretation follows context and identical phrases may give rise to several more or less idiomatic interpretations (2~3). This is impossible if its function is lexically fixed.

Table 118

2	Que la _x ballem	How we suffer!	(Espinal 2009)
3	Algunes dances _i , les _i ballarem a final de curs	We are going to dance some dances by the end of the course	
4	No sé pas com se les _i enginya...	I don’t know how (s)he manages...	=things _i
5	S’ho _i ha enginyat tan bé, que...	(S)he managed so well, that...	=it _i =situation _i

Conversely, the same phrase may use different clitics as appropriate to context whilst remaining idiomatic (4-5). (Un)idiomatic readings derive from each clitic’s [±individuated] status *in relationship to* discourse, not particular surface combinations. Moreover, ‘fossilization’ engendered by long-completed grammaticalization processes sits uneasily with the high synchronic productivity of such uses (Espinal 2009, also §1.3.2).⁹⁸ Such developments are only possible, if such clitics are recognized as regular syntactic elements with relatively fixed (if abstract) ‘readings’ (e.g. *ci*_[–individuated]=discourse-*here*), referencing a continuously developing shared encyclopaedic knowledge.

⁹⁸ Even clitics themselves may be productive. Mexican Spanish (Navarro 2005) has developed new uses: *la*=indeterminate/abstract object, *le*=abstract paths e.g. *pasarle* ‘go from one place to another’.

5.1.3 Range/Categories

Like SE (included for comparison), case is the primary divisor. *Ne* may substitute *part* of an item (ne_{PRT}) or reference the *whole* of which it forms (and remains) part (ne_{GEN}). Alternatively, it may reference the *place* whence it came (ne_{ABL}). That place may be abstract, representing previous *states* (discourse-*there/then*) left before entering discourse-*here/now*. These relationships may be oriented towards subject or object.

		NOM	OBL	DAT	ACC		
Function						Complement	
[- Individuated]	Domain Substitution	CI _{EXI}		LA _{ABS}		∅	Abstract
	Domain Reference	CI _{IMP}		CI _{IMP}		<i>a/con</i> +NP	
	Part Substitution	NE _{PRT}		NE _{PRT}		Partitive Article	Class
	Whole Reference	NE _{GEN}		NE _{GEN}		<i>di</i> +NP	
[+Individuated]	Spatial Reference	Subject-Orientation	CI _{LOC}	CI _{LOC}	Object-Orientation	<i>a/con</i> +NP	Place
	NE _{ABL}		(NE _{ABL})	<i>da</i> +NP			
	Internal Reference	SE _{ANT}		SE _{MID}	SE _{PASS}		
	External Reference	SE _{NOM}	OBL	SE _{DAT}	SE _{ACC}		
		Static		Dynamic			

CI_{LOC} references *places* at/to which the subject/object is/becomes present.⁹⁹ CI_{IMP} references SOAs as abstract domains. In some languages, it has ‘spread’ to represent the domain itself as subject, which is interpreted as ontological space and used in existentials (ci_{EXI}). In other languages, expletive subjects are depicted as possessing the item: ‘it_{EXPLETIVE} has many books’=‘There_{EXISTENTIAL} are many books’. These are generally represented by \emptyset_{NOM} clitics.¹⁰⁰

⁹⁹ Latin’s confusion of allative~locative continues in Romance.

¹⁰⁰ French as a non pro-drop language requires accompanying subject pronoun *il*.

LA_{ABS} can be seen as the object equivalent of *ci*_{EXI}, representing [–SPEC] objects impinging on the current state (*ci*_{IMP}). In Italian, it is also possible to distinguish discourse-*here/now*=*ci* from discourse-*there/now*=*vi*, although the difference is rarely observed.

Like DAT~OBL (Chapter 3), subject~object orientation is crucial. DAT relates to ACC, OBL relates to the event, and hence the subject. Similarly, locations may be subordinate to (and hence situate) objects (6) or event (7). These may coexist (8), because they modify different hosts; unlike two locations modifying the same object (9) or situation (10). Locative clitics equally appear in two positions. In (11), *y* situates the object, and appears under DAT. The subject may or may not be in the same place. In (12/13), *y* situates the subject and thereby the event. In (12), the subject must have arrived. *Y*_{OBL} indicates union with the place which is the existing discourse-*here*. (13) implies change of discourse-*here*; the subject was at X, but is now at Y where... Similar arguments can be made for ablative *en* (§5.2.2 and §5.3). Pescarini (2015, following Řezáč 2010) presents the order of French *en/y* as optional (14-15). In fact, such pairs illustrate different uses of *y*. (14) with *y* under DAT situates the object, whilst (15) with *y* under OBL, situates the subject.

Table 119

6	[SP]	Algunos chicos lo golpearon...en la cara	[lo _i golpearon [<i>e</i> _i [en la cara]] [Ø]]	
7		...en la clase	[lo _i golpearon [<i>e</i> _i [Ø]] [en la clase]]	
8		...en la cara en la clase	[lo _i golpearon [<i>e</i> _i [en la cara]] [en la clase]]	
9		...*en la cara en el ojo	Some boys hit him in the face, in the classroom	
10		...*en la clase en la escuela	(Sánchez Lopez 2007)	
11	[FR]	J' <y> vois [une chatte <dans le chambre?>]		
12		J'y arrive		
13		J' <y> vais <à Paris>		
14		Je te jure, j'en y ai vus trois	I swear, I saw 3 of them there	Object
15		Je te jure, j'y en ai vus trois	I swear, (while I was there), I saw 3 of them	Subject

Chapter 4 showed that SE displays static~dynamic oppositions: SE_{ANT}'s static current state resulting from a prior changes-of-state versus SE_{MID}/SE_{ACC}'s dynamic events changing the current state. Similarly for adverbial clitics. Subject-oriented *ne* references previous states, subject-oriented *ci* references static states, whilst object-oriented *ci* represents ever-changing discourse-*here*, or dynamic changes in object state with potential to change the current state.

5.1.4 Forms

Romance languages largely divide between those with (21-23), or without (24), adverbial clitics. Sardinian shows wide dialect/idiolect variation (Jones 1993:214-215). Unlike 'conservative' dialects (16), Campidanese has lost *bi*, and *n(ci)* (=in*ke*) is used for source, destination, and location (17), but many speakers replace *inke* with *inde* as source, freeing *inke* to express location/destination alone (18). Penello (2006) summarises dialect variation as (19-20). Examples such as (25) show clearly that *ne*_{PRT} and *ne*_{ABL} are distinct entities.

Table 120

		Partitive/Genitive	Source	Destination	Location/State
16	'Conservative' (Jones 1993)	inde	inke	bi	bi
17	Campidanese 1	inde	n(ci)	n(ci)	n(ci)
18	Campidanese 2	inde	inde	n(ci)	n(ci)
19	Baunese (Penello 2006)	inde	inde	(bi)	ince/je
20	Bittese/Ossi/Posadino	inde	(inde)	inke	bi
21	Italian	ne			ci
22	Catalan	en			hi
23	French	en			y
24	Spanish	Ø			Ø
		OF (di)	FROM (da)	TO (a)	AT (a)

	N	O	D	A	
25	bi _k	nke _j		nd' _i	at issitu [tres Ø _i] There _k came three _i (of them) out of there _j

5.1.5 Chapter Outline

In most languages, whether subject- or object-oriented, [−individuated] and [+individuated]

ci/ne have the same forms, whilst *ne*_{GEN} may be further confused with *ne*_{PRT}. It is in these uses that ‘idiomatic’ readings are found, and due to lack of recognition of these differences that lexicalization is invoked. The chapter proceeds by distinguishing each function/position for *ci* and *ne*. §5.5 onwards applies this understanding of available functions/interpretations to show that all ‘I-clitics’ are compositional. All that is required is recognition of both real and abstract referents (addressed by the movement from [+individuated] to [–individuated] within the interpreter) in relationships defined by case (and hence, in our model, position).

5.2 Object-Oriented Clitics

Whereas subject-oriented clitics operate in relation to the event (i.e. VP as a whole), Object-oriented clitics are within VP, as ACC (*ne*_{PRT}) or denoting a relationship to it (*ci/ne*_{ABL}).

5.2.1 *Ci*

The spatial proximal-distal continuum maps to grammatical person in pronominal domains. Most languages lost surface distinctions between locative pronouns e.g. French *y*, whilst others lost such clitics altogether e.g. Spanish. Early Italian personal pronouns *no(s)/vo(s)* ([±R]) were replaced by *ci/vi* which now exist independently of their spatial origins, including acting as reflexives of SE_{IMP} (§4.6.8-4.6.10).

Table 121

	Subj	Prep	Dat	Acc	Subj	Prep	Dat	Acc	Loc	Adv	Latin		[deictic]
I	io	me	me/i	mi	noi	noi	ci	ci	ci		hic	Proximal	+
II	tu	te	te/i	ti	voi	voi	vi	vi	vi	(ivi) ¹⁰¹	ibi	Medial	+
III	lui/lei	lui/lei	gli/e	lo/la	loro	loro	gli	li/le	(li)	(li)	illi ¹⁰²	Distal	±
	Singular				Plural				Adverbial				

101 Formal Italian also has a WP e.g. *ivi compreso* (Cardinaletti & Starke 1999:193). cf. French *ici* (<*i*+*ci*).
102 ILLI may also be dative.

The *here~there* distinction was largely lost from locative usage during the 1600's (Cortelazzo & Zolli 1999:1812). Modern Italian rarely shows differences between *ci/vi* in existential (26) or locative (27) usage. *Ci* as 'to/at here/there' is used in all circumstances, where *here~there* identifies the construal's situational focus rather than interlocutors. Formal/literary registers retain *vi*_{LOC} in situations expressing clear separation (28, Cardinaletti 2008:53). In (29, Russi 2008:58), *i pescatori ci vanno* construes a scene with the fishermen at the location (sentence subject's discourse-*here*), whereas in *vi sbarcheremmo* the sentence subjects are at a distance from the event (discourse-*there*). Presentational use of *ci/vi* (*ci presentativo*, Burzio 1986:126-132), is found from XIII^c. Use of *esservi* was previously significant, but now shows the same limited contrast as for locatives.¹⁰³ All ensuing examples use *ci*.

Table 122

26	V'/C' è modo e modo di farlo	There are better ways of doing it
27	Rimani qua/li? Sì, ci/vi rimango	Are you staying (t)here? Yes, I am staying (t)here
28	Gianni vi si oppone	Gianni opposed (himself) there
29	A:Ma nessuno va su quegli isolotti; sono isolotti sperduti. B:I pescatori ci _i vanno. Potremmo prendere una barca e una mezz'ora più tardi vi _i sbarcheremmo.	A:But nobody goes on those small islands; they are remote. B:Fishermen do go there _i (<i>ci</i>). We could get a boat and we would get there _i (<i>vi</i>) in half an hour.

Locative *ci* must reference discourse-salient location/situation/person/directions, hence (30~31, Maiden & Robustelli 2000:104-105)'s (un)acceptability. *Ci* pronominalizes PPs headed (32) by *a*, *in*, *su* ('topic' complements) and *con* (union, instrument, and material complements), often translated 'for/about/with it' (33-34). In each case, *ci* references participant coincidence with the SOA in spatial, temporal, or eventitive domains. *Ci* also corresponds to *assieme/insieme a* (39), where it may coexist with allatives. *Ci* indicates coincidence with the event, whereas allatives describe an event property. (32-34) reference objects within the transitive event, (37-42) reference the subject in relationship to that event.

¹⁰³ Many Sardinian dialects also retain proximate~distal interpretations (Bentley 2004:65, Loporcaro 1998:51).

Subject-oriented *ci* may also indicate stative relationships (42). This correlates with position.

The difference between static and dynamic *ci* for objects can be seen in (35-36); see (§6.3) for discussion.

Table 123

30	Guardo <u>sotto il tavolo</u> _i e <u>ci</u> _i troverai il fazzoletto	Look on the table and there you will find the napkin?	
31	A:Dov'è il fazzoletto? B:<*ce> lo troverai <là>	A:Where is the napkin? B:You will find it there	
32	[Sul/al tuo problema] _i <u>ci</u> _i ho pensato giornate intere	Topic	Object
33	[Con la lana avanzata] _i <u>ci</u> _i farò una sciarpa	Material	
34	[Con il cucchiaino] _i <u>ci</u> _i mangio la minestra di solito	Instrument	
35	me+ce+lo mette		
36	ce+me+lo mette		
37	[Con l'ombrello] _i <u>ci</u> _i uscirebbe anche in Giamaica	Union [-ANIM]	Subject
38	[Con Carlo] _i <u>ci</u> _i [esco spesso]	Company [+ANIM]	
39	[Assieme a Maria] _i <u>ci</u> _i [va sempre al cinema]		
40	<Ci> abita <a Roma>	He lives {there/in Rome}	
41	<Ci> va <a Roma>	He goes {there to Rome}	
42	C'è stato oggi	He is here, today	
43	La porta, <u>ci</u> ha dato un calcio	~gli ha data un calcio	C _{IMP}
44	Ce lo dico	=glielo dico	
45	Che ce la darestes voi vostra nipote?	Why, would you give her to him, your niece?	
46	A cosa/*dove Ø _i dedichi [il tuo tempo] _i ?		
47	<Ci> Ø _i dedichi molto tempo _i <al calcio>		

Ci_{IMP} substitutes *gli/le/loro* for inanimate recipients (43) and is often extended to reference persons (44),¹⁰⁴ when it may be used to 'breach' the PCC (45, Russi 2008:96). It represents the 'it/there' of current discussion. That *ci_{IMP}* is not truly locative, can be seen in *wh*-interrogatives where it is replaced by *che cosa*, not *dove* (46-47, Rigau 1982). None of these usages alternate with *vi* (Benincà 1988:177–78) which marks distal relationships.

¹⁰⁴ Berretta (1985a) delimits its use to specific regions and/or lower registers, however, Cordin & Calabrese (2001:576) describe its use in all regions, whilst Russi 2008:96-101) illustrates its widespread use among educated classes, including in writing.

5.2.2 *Ne*

As a [–DEF] clitic, *ne*_{ACC} pronominalizes NPs embedded under indefinite determiners (48), ‘partitive articles’ (49), and bare noun direct-objects in languages which admit them (51).¹⁰⁵ Like quantifiers, partitive articles (49, French *du/de la/des*, Italian *del/della/dei/delle*) act as weak indefinite determiners with null spell-out in the context of empty N(P)s (50), making it identical to use of bare nouns where the determiner is already null (51).¹⁰⁶ In these cases, *di* is not a preposition introducing PPs, but a [–DEF] case-marker. Since there is nothing to mark for empty DPs, the case-marker does not appear. This is confirmed by the fact that complements may retain other material (52).

When there is a specific class of items in local discourse, \emptyset _{DAT} may be interpreted as a weakly implied ‘of them’. When present, SE_{DAT} references subjects as possessors of *ne*_{ACC}’s [–DEF] object (54). Common in Old Italian, past-participle agreement is now largely restricted to pronominalized objects. Agreement with 1/2.ACC is optional (53, with no discernible semantic effect), required with 3.ACC (55). Since the *di* of partitive articles is a case-marker, [*dei libri*]_{ACC} causes agreement (50).

In its ‘genitive’ function, the direct-object is the noun (56, *un’altro*), whilst *ne*_{GEN} substitutes the *di*-phrase (a true prepositional phrase), referencing the class/set of items from which the nominal originates. In these cases, agreement with the past participle is not allowed, since the accusative has not been pronominalized. In many cases, nominal and adjectival readings are available, in others, presence of datives (54-55), or past-participle agreement (57-58) determine a particular reading.

¹⁰⁵ Italian also allows fractional nouns, where verbal agreement is with the quantifying nominal, not the *de*-phrase DP: *Ho comprato delle mele_i e ne_i ho mangiata la metà_i.*

¹⁰⁶ Longobardi (1994) for the presence of null D in argumental bare nouns.

Table 124

	D	A			Did you bring any books?	Italian
48		ne _i		[due/molti/alcuni [Ø _i]] _A	Yes, I have brought...two/many/a few	
49		<ne>		<[dei _{PRT} libri] _A >	...some _{PRT.PL} books	
50	Si,	Ø _j	ho portati	[di [Ø _i]] _A		
51		ne _i		[[Ø _i]] _A	...some (bare noun)	
52		ne _i		[di belle [Ø _i]]	...some of the good ones	
53	Non	me _i	ha visto/a	e _i	He didn't see me	
54		me	<ne>	compro	[una macchina]	I bought myself a car
55		me	la _i	sono comprata	e _i	I bought it for myself
56		<ne _j >	Ø _i	prende	un'altro _i <dei libri _i >	He takes another {of them/of the books}
57	Di mele _j ,		ne	ha mangati	due chili e _j	Of apples, he ate...[some 2 kilos] _{ACC}
58		ne _j	Ø _i	ha mangato	[due chili _i]	...[2 kilos] _{ACC} of them _{GEN}

French follows the same pattern (59-62) including past-participle agreement¹⁰⁷ with cliticized (63) and *wh*-fronted (64-65) direct-objects, but not indirect-objects. As a partitive case-marker, *direct-de* never takes wide scope over coordinated phrases (66),¹⁰⁸ whilst as a preposition introducing an independent phrase, *indirect-de* may (67). *Y* and *en_{GEN}* pronominalize indirect *à/de*-PPs introducing undifferentiated notions equivalent to *cela*. Since *penser* is not an indirect-transitive, human dependants *à mes frères/à eux* (68) cannot be indirect-objects, and hence cannot cliticize as *leur_{DAT}*. *Y* represents not *à eux*, but *à cela*. Semantically, *eux* regards the brothers as individuals, whilst *y* views them as an undifferentiated set. Similarly, *en_{GEN}* replaces *de cela* (69).

Neuter possessors display *en_{GEN}* (70~71). Although Italian does not use clitics to represent inalienable possession, *ne*-extraction still requires possessive DAT/ACC relationships (Longobardi 1991:59). Whilst (72) admits two readings, (74) only accepts experiencer readings, as do cases of *ne*-extraction (75).

107 French past-participle agreement is unstable. It is largely orthographic, unmarked phonologically for *-er* verbs (the largest category). It may surface orally with a small set of irregulars e.g. *dire*, marking gender, but not number: *dit(s)* [di]~*dite(s)* [dit], but is generally poorly respected (Goosse 2000:126).

108 A single *à/de* may scope over VPs containing coordinated Vs (Abeillé & Godard 1997).

Table 125

59	J'<en> ai apporté deux/beaucoup/quelques-uns <livres>	I have brought two/many/some books	French
60	J'<en> ai apporté <des livres>	I have brought some books	
61	J'en _i ai apporté [de Ø _i] _{ACC}	I have brought some	
62	J'en _i ai apporté [de bons [livres _i]]	I have brought some of the good ones	
63	Les maisons, je les ai repeintes	I repainted the houses	
64	Quelles maisons avez-vous repeintes?	Which houses did you repaint?	
65	Les maisons que vous avez repeintes	The houses you repainted	
66	Il y avait sur la table beaucoup de pain et *(de) vin	There was a lot of bread and wine on the table	
67	J'ai besoin de [cette farine et cette levure]	I need this flour and baking powder.	
68	Mes frères, je <*leur/°y> pense souvent <à eux>	I often think about my brothers/them.	
69	Mes deux filles, je <*leur/°en> dépends <d'elles>	I depend on my two daughters/them	
70	M <lui> Ø _i a cassé [le bras _i [<de P _i >]]	[+ANIM] M. broke P.'s arm	
71	M <en> Ø _i a déchiré [la page _i [<du livre>]]	[-ANIM] M. tore the page of the book/it	
		I remembered... the desire of G.	Italian
72	Ø Ø _i ho ricordato [il desiderio _i [di G] _{ADJ}]	...=X (usually subject)'s desire for G.	
73	Ø Ø _i ho ricordato [il desiderio _i [di G] _{GEN}]	...=G.'s desire (for something)	
74	Ø _j Ø _i ho ricordato [il suo _j desiderio _i e _j]	...his desire	
75	Ne _j Ø _i ho ricordato [il e _j desiderio _i e _j]		

Gross (1968) observes that in (76-77) and (79-80) each argument may pronominalize separately, they cannot co-occur (78, 81). This may be a 3-3-rule for some speakers, but cases occur. Jones (1996:254) labels *y+en* (82) and *en+en* (83) as 'literary' and 'atypical'. Non-standard varieties with different D/A swapping rules (§6.10.3), show *en+y* (84, Ayres-Bennett 2004:209). Another confusion arises in French combinations with personal pronouns. In (85), clitics appear as expected, but in (86) they swap due to relative weight (§6.10.3).

Table 126

	S _H	D	A	X		
76	Je	Ø _i	<en _i >		vois un <chat _i > dans la chambre _i	I see {a cat/one _i } in the room _i French
77	J'	<y> _i	Ø _i		vois un chat _i <dans la chambre> _i	I see a cat _i {there/in the room} _i
78	*J'	y _i	en _i		vois un e _i e _j	*I see one _i there _i
79	Il	Ø _i	<en _i >		remplit un <verre> _i de ce vin _i	He fills {a glass/one _i } with this wine _i
80	Il	<en _i >	Ø _i		remplit un verre _i <de ce vin> _i	He fills a glass _i {with this wine/of it _i }
81	*Il	en _i	en _i		remplit un e _i e _j	*He fills one _i with it _i
82	Il	y _i	en _i		a acheté deux e _i e _j	He bought some two (from) there
83	Il	en _i	en _i		a acheté deux e _i e _j	...of them
84	%J'	■	en _i	y	ajouterais régulièrement	I would add some to it regularly
85	M	lui _i ⁺	en _i ⁺		donnera	I will give some _i to him _i
86	M	■	l _i ⁺	en _i ⁺	informera	I will inform him _i of it _i

Use of *ne* to reference object spatial origins, was common in Old Italian with *ne+lo* still available in some varieties (87-88, Lepschy & Lepschy 1984:212). In Modern Italian, it is infrequent, only occurring in isolation. Object-oriented *ne*_{ABL} is more common in Catalan e.g. (91, Cortés & Gavarró 1997), where the sense of ‘from within’ (89) or even static ‘in’ (90) requires *dentro* in Italian. In Italian/Catalan, combinations with *ne*_{ABL} are generally expressed by locatives: Italian *ne*⁺+*ne*⁺→*ci*⁺+*ne*⁺, Catalan *hi*⁺+*en*→*n*’*hi*⁺ (§6.4.2).

Table 127

87	Ne lo trasse	He pulled it out from there	Italian
88	Ne lo liberava	He was freeing him from it	
89	L’ho preso dentro il cassetto	I took it from (out of) the drawer	
90	L’ho trovato dentro il sacco	I found it {in(side)/(with)in} the bag	
91	<En _{ABL} > trec l’abric <de l’armari>	I take the coat out of the cupboard/it	Catalan

5.2.3 Object-Clitic/Functions

(92) summarises the uses of object-oriented clitics discussed above. Similar patterns are found in other languages and dialects, often with different swapping patterns and 3-3-rules, as discussed in Chapter 6. It is clear that the number of functions does not match the number of forms. In our opinion, it is the attempt to treat them as one-to-one correspondences that leads to confusion and invocation of lexicalization.

Table 128

92	1	2	3F	3M	3N		LOC	ABL	3-3
Italian	mi	ti	le	gli	ne _{GEN}	ci _{IMP}	ci _{LOC}	ne _{ABL}	glie ₃₋₃
French	me	te	lui		en _{GEN}	y _{IMP}	y _{LOC}	en _{ABL}	lui ₃₋₃
Catalan	me	te	li		en _{GEN}	hi _{IMP}	hi _{LOC}	en _{ABL}	hi ₃₋₃
Spanish	me	te	le			Ø			se ₃₋₃
DAT	-E					+E			

5.3 Subject-Oriented *Ne*

Under the unaccusativity hypothesis, both the possibility of *ne*-extraction from post-verbal unaccusative subjects (93) and its impossibility with unergatives (94) derives from the assumption that *ne* is an object-only clitic i.e. unaccusative subjects are ‘deep objects’. In fact, subject *ne*-extraction from unergatives is wide-spread (Italian, Lonzi 1986; French, Hulk 1989; Catalan, Cortés & Gavarró 1997) showing that this assumption is incorrect. Conversely, *ne*-extraction is impossible from animate subjects with certain unaccusatives (Lonzi 1986:114). Unaccusativity~unergativity cannot determine *ne*-extraction’s availability.

Table 129

93	[IT]	Ne _i arriveranno [molti <i>e_i</i>]	Many will arrive
94		*Ne _i telefoneranno [molti <i>e_i</i>]	Many will telephone
95	[FR]	Il en arrive deux	Two of them arrive
96	[CA]	N’han arribat 22.511	22,511 have arrived
97		De 1.200 habitants en van morir 110	Out of 1200 inhabitants 110 died
98	[FR]	Pourtant il en volait encore en 1978	However some were still flying in 1978
99	[CA]	En van correr més de 40	More than 40 ran
100	[IT]	Su 13 mezzi acquistati ne camminano solo 6	Out of 13 trams only 6 work
101		Tre di loro sono stati <u>uccisi</u>	Three of them have been killed
102		Ne sono stati uccisi tre	
103	[CA]	Malauradament algunes s’han <u>perdut</u>	Unfortunately some have been lost
104		Se n’han perdut algunes. Les que s’han conservat...	Some have been lost. Those that remain...

In addition to passives (102), anticausatives (104), and other prototypical presentational¹⁰⁹ intransitives (95-97), *ne/en* occurs with other verbs when used with presentational import (98-100). Conversely, extraction is unavailable from all such verbs when focused i.e. with identificational information structure. Thus (101, 103), but not (102,104), allow stress on the verb. French subject *ne*-cliticization is overtly restricted to expletive-inversion, but restriction to presentational structures is also true of Catalan/Italian; merely less apparent without overt *il*_{EXPLETIVE}. Transitive direct-objects (without marked intonation) and (expletive) associates

¹⁰⁹ ‘Prototypical presentationals’ include: presentational verbs in the strict sense e.g. *arrive*, *appear*; verbs that may be used presentationally e.g. *die* (a-b); passives with indefinite post-verbal subjects.

(a) [SP] Murió mucha gente (Presentational) Many people died (=There were many deaths)

(b) [SP] Mucha gente murió (Non-presentational) Many people died (=Many individuals suffered death)

represent the same presentational information structure, withholding focus from the verb. Thus, the apparent link between unaccusatives/passives and *ne*-cliticization reflects natural presentational capabilities of *some* unaccusatives, and passives in general. Parallel syntax is unnecessary. For similar arguments, see Lonzi 1986; Levin & Hovav 1995:276-7; and Mackenzie 2006 (from which many of the examples are drawn).

Bentley (2004:237-8) argues that Italian *ne*-extraction does make focus-based unergatives~unaccusatives distinctions: subject *ne*-extraction is compatible with wide and narrow quantifier focus with unaccusatives, but only wide focus with unergatives, as shown by its unacceptability in interrogative structures and their replies from unaccusatives (105), but not unergatives (106). French *ne*-cliticization, however, *can* appear under narrow focus with unaccusatives, passives and unergatives (107-109). Italian also admits cases like (106), given suitable context and/or non-agentive activity verbs (110-111). Agentive activity (*camminare*'s default sense) semantics clash with presentational contexts required for *ne*-extraction. Presentational occurrences of such verbs create weak existential interpretations back-grounding verbal agentivity (112) in contrast to the default 'identificational' information structure applied out-of-context (Pinto 1997:21-22). (106)'s deviancy derives from lack of suitable context leading to agentive readings. Contra Bentley (2004), being interrogative or having narrow focus are irrelevant. Supporting context is sufficient to ameliorate such deviancies (110), whilst with non-agentive unergatives (111), explicit contextualization may not even be required.

Catalan (Cortés & Gavarró 1997) confirms subject *ne*-extraction's relationship to agentivity and/or information structure. *Menjar* may be used (in)transitively. As an unergative, external

arguments may undergo *en*-cliticization (113). As a transitive, themes may do so (114), but agents are blocked (115). The same results obtain with quantifiers modifying *en* (116-117).

Table 130

105	Quanti ne muoiono/nascono/arrivano?	How many (of them) die/are born/arrive?
106	??Quanti ne camminano?	How many (of them) walk?
107	Q:Combien en est-il resté en France? A:Il en est resté moins de quatre mille	How many remained in France? Less than four thousand remained
108	Q:Combien en a-t-il été produit? A:Il en a été produit des centaines	How many were produced? Hundreds were produced
109	Q:Combien en vole-t-il au dessus de la ville? A:Il en vole trois par jour	How many fly over the town? Three fly over per day
110	Lawyer: Quanti aerei partecipavano a quella... Witness: Eh, non mi ricordo Lawyer: Generalmente quanti ne partecipano?	How many aircraft were participating in that... I don't remember Generally how many participate?
111	Quanti ne funzioneranno?	How many of them will be working?
112	Nell'amministrazione lavorano numerose donne, generalmente mal retribuite	In public administration many women work, generally poorly paid

As indicated by quantifier position, extraction from pre-verbal position is ungrammatical even for inherently presentational unaccusatives/passives. The verb's external argument (merged at SPEC,vP) may raise *iff* it is agentive to SPEC,IP, where it is 'higher' than the clitic position which *ne* targets. Non-extraction reflects scope, *not* subject~deep-object, or unaccusative~ergative. Scope is a product of presentational~identificational information structure, itself reflecting subject (non)agentivity.

Table 131

113	Quantes persones van menjar a la cuina? How many people ate in the kitchen?	–N'hi van menjar sis Six of them ate there	Catalan
114	Quantes pomes van menjar? How many apples did they eat?	–En van menjar moltes They ate many of them	
115	Quantes persones han menjat gelats, avui? How many people have eaten ice cream today?	–*N'han menjat tots gelats, avui All of them have eaten ice cream today	
116	En _i vindran tres _i massa tard a la reunió Tres _i *en _i vindran massa tard a la reunió *En vindran tard a la reunió tres Three of them will come late to the meeting	117 En seran convidats molts, a la revella *Molts en seran convidats a la revella *En seran convidats a la revella molts Many of them will be invited to the party	

Ne is *not* an object-only clitic, but may represent subject-oriented participants extracted from post-verbal associates in presentational clauses. This is supported by the development of Romance, where extension from object- to subject-oriented partitives is a necessary pre-requisite for development of partitive-articles (e.g. French *du*), pre-dating the rise of object- and subject-oriented *ne*_{GEN} (Carlier 2007).

5.3.1 *Ne*_{NOM}~*Ne*_{OBL}

Contra many earlier works, post-verbal position cannot be assigned object θ -role (=deep-object). Chomsky (1995:274) notes that Italian post-verbal unaccusative subjects behave as pre-verbal subjects with respect to control; subjects, but not objects, are sufficiently ‘high’ to c-command into adjunct clauses (118). This is true of all Romance pro-drop languages, for unaccusatives (119), and unergatives with agentive (121) or theme (122) subject. Its impossibility in semantically identical non pro-drop French (120) implies that the phenomenon is structural in origin. In Chomsky (1995), pro-drop control patterns derived from covert raising of subject features to high pre-verbal positions unavailable in French, from which the fixed singular verb derives. Such feature movement hypotheses, however, have been abandoned. Under the minimalist program, subjects remaining *in situ* take nominative case entering into LDA with T which c-commands it (Chomsky 2000:122-3).

Table 132

118	[IT]	Sono entrati tre uomini _i [senza <i>pro</i> _i indentificarsi]	}	3 men entered without identifying themselves
119	[SP]	Entraron tres hombres _i [sin <i>pro</i> _i identificarse]		
120	[FR]	*Il est arrivé trois hommes [sans <i>pro</i> _i s’identifier]		
121	[SP]	Gritaron tres hombres _i [sin <i>pro</i> _i identificarse]		3 men shouted without...
122	[SP]	Ha muerto mucha gente [sin <i>pro</i> _i hacer un testamento]		Many people have died without making a will

Following this approach, *dei ospiti* is the subject in post- (123) and pre-verbal (124) positions, pronominalized as *ne*_{NOM} (125-126). In (123/125), *S_H* is empty, because the subject is present; overtly (*S_L*) or as *ne*_{NOM}. When extracted from the clause (126), subject arguments must be filled. If this were *ne*_{OBL}, (125-126)'s subject would default to *they* contrary to meaning. A weakly-implied *of them* derives from \emptyset _{OBL} (=indirect-subject) related to NOM (=direct-subject), just as \emptyset _{DAT} (=indirect-object) may imply arguments for ACC (=direct-object).¹¹⁰

In (127), *tre* modifies the subject pronominalized as *ne*_{NOM}. It may be focus-fronted (128), where its adjectival status is intonationally highlighted, presaging its contrast with *dieci*. *Ne*_{NOM} references the discourse topic. In both cases, \emptyset _{OBL} implies a weak *of them*. Without contextual information and under normal intonation, the adjective may be interpreted as evidence of nominal ellipsis (129), possibly raised to *S_H* (130). In such cases, *ne*_{NOM} would double [_N \emptyset] and so is unacceptable. The weak *of them* in (129,130) may be made explicit (131,132). Cardinaletti & Giusti (2006:114) claim that *ne* must be absent with non-anaphoric [_N \emptyset], however, many Italian speakers require *ne* in all cases (Lepschy 1989). This may be *ne*_{OBL} referencing generic types/elements understood from discourse (Corblin 1995), or surface-identical *ne*_{NOM} (127) through *ad sensum* reference.

In (133), *SE_{ANT}* is the nominative reflexive of subject *vasi*; there is no anaphoric reference since the information is new, and hence no implied *of them*. In (134), the ellipsed noun requires that a class referent be found from discourse i.e. \emptyset _i=*examples of what is under discussion*, hence introducing a weak *of them* (\emptyset _{OBL}). That class may be made explicit, via PP or clitic (136). Note that the alternative reading for (134) is not available (135), since *SE_{ANT}*

¹¹⁰ Use of *some* occasionally makes English translations awkward, but has the benefit of clearly separating nominative (direct) *some* from oblique (indirect) *of it/them*.

(required to make *rompere* intransitive) already occupies NOM. Similarly, personal SE_{ANT} (137). Post-copular subjects are often read as weak existentials (138) with no implied *ne*_{OBL} (in contrast to 139), but where *ne*_{NOM} references that wider element. Existential readings can be made explicit by nominative *ci*_{EXI} (140, §5.4.3). Hence, there can be no ‘there are some three’, only ‘there are three of them’.

Table 133

	Topic	S _H	N	O	V	S _L	
123		Ø _i	Ø _i	Ø	sono rimasti	[_{DP} dei ospiti _i]	Some guests remained
124		[_{DP} dei ospiti _i]	Ø _i	Ø		[_{DP} e _i]	
125	(X _i)	Ø _i	ne _i	Ø _j	sono rimasti	[_{DP} e _i]	Some _i remained
126	[_{DP} dei ospiti _i],						Of the guests _j , some _i remained
127	X _i	Ø _i	ne _i	Ø _j	sono arrivati	[_{NP} tre [_N e _i]]	Some 3 arrived
128	X _i , No, TRE _k ,					[_{NP} e _k [_N e _i]] non 10	No some 3 arrived, not 10
129	X _i	Ø _i	Ø _i	Ø _i	sono arrivati	[_{NP} tre [_N Ø _i]]	3 (TOPIC _i) arrived
130	X _i	[_{NP} tre [_N Ø _i]]	Ø _i	Ø _i		[e]	
131	X _j	Ø _i	Ø _i	ne _j		[_{NP} tre [_N Ø _i]]	3 (of them _j) arrived
132	X _j	[_{NP} tre [_N Ø _i]]	Ø _i	ne _j		[e]	
133		Ø _i	se _i	Ø	sono rotti	[tre vasi _i]	3 vases broke
134	(X _i)			Ø _i		[tre Ø _i]	3 Ø (of them) broke
135	(X _i)	Ø _i	*ne _i	Ø _j		[tre Ø _i]	*Some 3 (of them) broke
136		Ø _i	se _i	<ne _j >		[tre Ø _i <dei vasi _j >]	3 Ø {of them/the vases} broke
137	(X _i)	Ø _i	se	ne	sono perduti	sette	7 of them were lost
138		Ø _i	ne _i	Ø	sono morti	[_{NP} tre [_N Ø _i]]	There have been 3 deaths
139	(X _j)	Ø _i	Ø _i	ne _j			3 of them _j died
140	(X _i)	Ø _i	ce	ne _j	sono	[_{NP} tre [_N e _i]]	There are 3 _i of them _j
141	[della rivolta _j],	[una foto _i]	Ø _i	ne _j	fu	la causa	A picture was the cause <u>thereof</u>
142	[de ce livre _j],	[le premier chapitre _i]	Ø _i	en/y	est	intéressant	The 1 st chapter <u>thereof/therein</u> is interesting

Ne-extraction to OBL from raised subjects is rare (Belletti & Rizzi 1981:120; Burzio 1986:30-31) but sometimes found (141, Moro 1997:60). Pollock (1998:307) notes that *en/y*-cliticization is acceptable for some French speakers where it would be inadmissible in Italian (142). In both cases, referents must be readily accessible from context. *Ne*-extraction should not be available from raised subjects, since they are already higher than NOM/OBL. In such

cases, *ne*_{OBL} references *not* the subject/associate or dependent PP e.g. *la causa della rivolta*, but implicit/explicit dislocated topics (made explicit in 141-142). In this sense, they are no different from any clitic pronominalizing dislocated referents. Acceptability depends upon context, and language-specific restrictions upon topicalization.

Context is everything. Spanish (143-144) introduce new topics (*elettori/persone*), while (145-146) are discourse-dependent. (143) may take contrastive (143a, *ne*_{NOM}+*ad sensum* reference) or neutral (143b, *ne*_{OBL}+part-whole reference) readings. The difference is slight. In (144), however, the two sets of people are logically disjoint i.e. 26,000 cannot be part of the whole (*of them*) represented by 7,500 which is the only anaphoric referent available under locality. In (144)'s first clause, *ne*_{NOM} is inappropriate since it would double the explicit subject *persone_i*. In the second clause, *ne*_{NOM} is the pronominalization of *persone_j*, a different set from *persone_i*. Neither clause requires reference to any prior set, such that *ne*_{OBL} is not required. In (145-146), *ne* in the first clause highlights contrast between the two groups taken to be drawn from specific (145, *ne*_{OBL}) or generic (146, *ne*_{NOM}) anaphorically referenced groups. (147-148) are matching examples from Catalan.

Table 134

	Topic	Statement	
143	Su 721 elettori	a. <i>ne</i> _{NOM} b. <i>ne</i> _{OBL}	hanno votato 635 a. Out of [721 voters] _i , some 635 _j voted b. Out of [721 voters] _i , 635 _j of them _i voted
144		Al CNR lavorano [7.500 persone _i], mentre al CNRS <i>ne_j</i> lavorano [26.000 persone _j]	[7,500 people _i] work at the CNR, while [some _j 26,000] work at CNRS
145	Dei X _x ,	Al CNR <i>ne_x</i> lavorano [7.500 Ø _i], mentre	Of the X, 7,500 _i of them _x work at CNR,
146	Ø _x ,	al CNRS <i>ne_x</i> lavorano [26.000 Ø _j]	while 26,000 _j of them _x work at CNRS (Of people _x ,) 7,500 _i work at CNR, while 26,000 _j work at CNRS
147	Sobre 1.622 persones,	en voten 601	Out of 1,622 people, 601 vote
148		Som en plantilla 50 persones, però en treballen moltes més cobrint baixes	We are a basic team of 50 people, but many more work cover absences

Catalan provides further support for ne_{NOM} . Some Catalan speakers admit extraction without quantifiers (149, Fabra 1956) i.e. without overt source. Since DPs containing ellipsed nouns cannot be postulated without a quantifier, (149) must be a pro-drop subject represented by $ne_{NOM}=some$. If ne were OBL, (149-150)’s subject would default to *they*, i.e. ‘**they of them sleep*’. Indeed, a weak *of them* (whole) can only be implied if there is a subject *some* (part) from which to reference. In the presence of overt or clitic class reference, readings with definite subjects are required (151). When the class reference is topicalized, OBL continues to reference it, and NOM continues to reference the subject as definite pro-drop subject (152) or ne_{NOM} (153). In French expletive-inversion, subject *il* appears under S_H (154) matching \emptyset_{NOM} , hence only one reading is available, although translation as ‘some two arrived’ is common.

Table 135

	Topic	S_H	N	O	V	S_L	
149	X_i	\emptyset_i	en_i	\emptyset_x	dormen	\emptyset_i	Some sleep
150						[tres e_i]	Some three sleep
151	[Dei X_j],	\emptyset_i	\emptyset_i	<en $_j$ >	dormen	[tres \emptyset_i <dei X_j >]	Three, of {the X /them $_j$ } sleep
152			\emptyset_i	\emptyset_j		[tres \emptyset_i [e_j]]	Of the X ,...three sleep
153			en_i	\emptyset_j		[tres e_i [e_j]]	...some three sleep
154	[Des X_j],	Il_i	\emptyset_i	en_j	arrive	deux	Two of them arrive

Subject-oriented Class substitution (NOM) and reference (OBL) are ‘blocked’ by the presence of objects and, therefore, only available with intransitives or presentational transitives. Since they are mutually exclusive with object-oriented substitution (ACC) and reference (DAT), the two pairs have been treated as the same items generating the complexities of ‘deep-objects’. This analysis follows modern theory in treating these arguments as (in)direct-subjects and hence able to enter into LDA with higher functional positions: S_H /NOM/OBL.

5.3.2 *ne*_{ABL}

Separation of *ne*_{ABL} is justified by its different etymology and form in Sardinian (§5.1.4), the nature of its referent, and syntactic behaviour. Such uses are subject-oriented and only found with intransitives, passives or presentational transitives (155-156). In resultative passives, *ne*_{ABL} substitutes *da*+NP indicating the source/cause of resulting physical/mental states (157-158). Its use, often incorrectly treated as lexicalized, is exemplified in §5.5.6.

Table 136

155	[IT] Si avvicinò le zampe e poi se ne allontanò	It approached the harbour and then went away from it
156	[FR] Il n'<en> est jamais sorti [pp <de là(-bas)>]	He has never come out from there
157	[IT] I tulipani <ne> furono distrutti <dal vento>	The tulips were destroyed by the wind
158	Quando noto una contraddittorietà, ne resto turbato	When I see a discrepancy, I am disturbed by it

5.4 Subject-Oriented *ci*

Subject-oriented *ci*_{OBL} as contrasted with object-oriented *ci*_{DAT} was introduced in §5.2.1. This section provides evidence for nominative *ci*, which developed from *ci*_{OBL} in some languages in order to express existentiality (*ci*_{EXI}). Italian also has a form equivalent to lower clitic-field *ci*_{IMP}, permitting a range of additional readings not found in other languages. The section ends with consideration of the different ranges of existential clitics found across Romance.

5.4.1 Existentials

Cross-linguistically, existence is equated with placement in abstract space (Lyons 1968). Freeze (1992) and Moro (1998) *i.a.* treat existential and locative constructions as equivalents, however, this cannot explain the breadth of synchronic/diachronic variation found across Romance (§5.4.2). We follow McNally (1992), Zamparelli (2000), Remberger (2009), Cornilescu (2009), *i.a.* in maintaining a fundamental locative~existential distinction.¹¹¹

¹¹¹ Francez (2007) and McNally (2011) for literature overview.

Locative (159) and existential (160-161) sentences represent different perspectives (Partee & Borschev 2002, 2007), where one element is highlighted and the rest is predicated of it. PP_{LOC} is obligatory in locative (159), but optional in existential (160-161) sentences (Zamparelli 1998; Hazout 2004; *i.a.*). (160) centres upon the abstract space of existence, asserting a content property. Further locations (161) intersect with abstract space making the context more specific, without changing the nature of that assertion. A similar shift in perspective is seen with ‘atmospheric predicates’. (162) has a referential subject as its perspectival centre (which happens to be a location) about which properties are asserted. In (163), the perspectival centre is expletive *it*. Without further context, default ‘atmospheric’ readings are inferred, relating to *here-and-now*. The space over which this property holds may be further specified by additional locatives (e.g. *in the room*). Whilst English distinguishes *it~there*, other Germanic languages use *it* for both functions (164, German), whilst African American English alternates *it* with arbitrary *they* (165, Green 2002:80). In many languages e.g. Hebrew (Hazout 2004:413) and Romanian (§5.4.4), the locative centre is covert, although there are clear linguistic clues which signal its presence.

Table 137

159	Many girls are *(in the room/there)	BE (PROPERTY, LOC _[+SPEC])
160	There were many girls	BE (LOC _[-SPEC] , PROPERTY)
161	There are many girls in the room	BE ((LOC _[-SPEC] , PROPERTY), LOC _[+SPEC])
		LOC = LOC _[-SPEC] \cap LOC _[+SPEC]
162	The room is cold	
163	It is cold (in the room/here)	
164	Es ist ein Buch auf dem Tisch	There (lit. it) is a book on the table
165	{It/Dey} {got/have} some coffee in the kitchen	There is some coffee in the kitchen

Pragmatically, existential sentences “introduce the NP referent into the discourse world of the interlocutors by asserting its PRESENCE in a given location” (Lambrecht 1994:179).

Existentials presuppose locations, hence (166) is infelicitous in out-of-the-blue contexts, but acceptable in (167). Access to prior locations is DP-dependent. Thus, (168) is acceptable because its referent naturally accesses ontological space, however, *cockroaches* (166) presuppose [+SPEC] locations. Existential DPs take focus and must be hearer-new.¹¹² English only accepts [−SPEC] referents, except in special interpretations (e.g. lists, Milsark 1974). In addition, Italian has a similar construction allowing [+SPEC] referents (§5.4.3).

Table 138

166	?There are cockroaches.
167	Don't go into the kitchen. There are cockroaches.
168	There is a God.

5.4.2 Romance Existentials

Classical Latin employed ESSE, with HABERE (taking nominative or accusative pivots) appearing in late Latin (Cennamo 2011). Early Italo-Romance shows existential constructions similar to locatives and possessives (Ciconte 2010). Existential clitics are Romance innovations.¹¹³ Cruschina (2014)'s survey of 115 Italo-Romance dialects, found that languages either possess identical locative and existential clitics, or neither.¹¹⁴ Proforms are missing in Romanian, Ladin, Friulian, Romansch, some Venetian and southern Italo-Romance dialects, European/Brazilian Portuguese. Spanish, Galician, and Asturian show lexicalized proform -y solely in the present tense verb.

Table 139

ESSERE	Romanian, Italian, Corsican, Friulian, Romansch, Ladin, and many Italo-Romance dialects
STARE	Some southern Italo-Romance dialects
TENERE	Brazilian Portuguese
HABERE	Spanish, Asturian, Galician, European Portuguese, French, Catalan, some Salentino/Calabrian dialects

¹¹² The *Novelty Condition* of McNally (1992).

¹¹³ For etymologies: Rohlf (1969:899), Maiden (1995:167), Blasco Ferrer (2003) and Benincà (2007).

¹¹⁴ Some Calabrian dialects, otherwise lacking locative clitics, have borrowed whole existential *ci* constructions from Italian (Sorrenti, *in prep.*).

In southern Italo-Romance, HAVE-existential pivots never show agreement (169, Martano), being syntactically marked as direct-object by displaying prepositional-*a* following dialect-dependent rules (Bentley *et al.* 2013 for examples), whilst dislocated (170) or resumed (171) pivots display accusative clitics. Direct-object status of HAVE-existential pivots is also claimed for Spanish (173, Suñer 1982) and Catalan (172, Rigau 1994, 1997). Spanish HABERE existentials never exhibit personal-*a*, however, they do show direct-object (partitive given their [–DEF] referents) resumptive clitics (173, Leonetti 2004). The 3.SG verb of HAVE-existentials points to interpretations as impersonal constructions with object pivot and null subjects surfacing as expletive pronouns in non pro-drop languages like French (174, (Giurgea 2012). Accompanying locative clitics are exactly what they seem. Several central/southern Italo-Romance dialects (Ledgeway 2008, 2009:ch.16) employ STARE (175, Macerata, Marche), where *ci* and PP_{LOC} are mutually exclusive (176) and agreement is shown when distinct 3.SG~3.PL forms are available: Macerata only has *sta*. With contextually determined indefinite pivots, the same surface sequence may take existential readings.

Table 140

169	Intra lu cassettu, li sciucamani, non l'ave	In the drawer, there are no towels	Martano
170	Non l' _{ACC} ave, soruta, intra l'ufficiu	Your sister isn't there, in the office	
171	T' _{ACC} ave a la festa? –Sì, m' _{ACC} ave	–Will you be (there) at the party? –Yes, I will	
172	A la reunió hi havia el president	The president was at the meeting	Catalan
173	–Hay brujas? –Sì, las hay	–Are there witches? –Yes, there are	Spanish
174	Il y en avait deux	There were two of them	French
175	Le pantofole sta sotto lu lettu	The slippers are under the bed	
176	Ce sta le pantofole, sotto lu lettu	There are the slippers,.../The slippers are there,...	

Existential sentences without pro-forms are attested throughout Old Romance. Ciconte (2009, and examples therein) illustrates existential and locative sentence development from early Tuscan to Modern Italian: [–existential,–locative] readings with no clitics (177), locative

readings with adjunct (178) or clitic (179), and the impossibility of two clitics (183), remain constant throughout. What changes is the means of indicating existentiality and its relationship to locative expressions.

Table 141

	Reading		N	O		Tuscan		Mod. Italian	Reading	
	Exi	Loc				XIV ^c	XVI ^c		Exi	Loc
177	–	–	Ø	Ø		✓	✓	✓	–	–
178	–	+	Ø	Ø _i	L _i	✓	✓	✓	–	+
179	–	+	Ø	vi _i	e _i	✓	✓	✓	–	+
180	+	–	Ø	Ø		✓		✗		
181	+	–	vi _j	Ø		✓	✓	✓	+	–
182			vi _j	Ø _i	L _i	✗		✓	+	+
183	+	+	vi _j	vi _i	e _i	✗	✗	✗	+	+

Early Tuscan showed complementary distribution between PP_{LOC} and existential readings with both overt (181) and covert (180) existential-marker. Presence of PP_{LOC} debarred existential readings in covert existentials, and was illicit with overt existential-markers. During XVI^c,¹¹⁵ use of covert existentials declined, so that such sentences today may only take non-existential readings (177). Increase in overt existential clitics (181) was accompanied by co-occurrence with PP_{LOC} (182), and the modern situation where *ci* is required for existential readings, and *ci*_{EXI}+PP_{LOC} is acceptable. All modern languages admit co-occurrence of locatives and (c)overt existential-markers, indicating that existentials are *not* locatives.

¹¹⁵ Similar developments are found in Roman, Campanian and Sicilian during XIV^c-XV^c.

5.4.3 Italian

Italian use of *ci* in this area is multi-faceted. We analyze the NOM~OBL and [+individuated]~[−individuated] distinctions in terms of four categories as illustrated and contrasted in Table 142 and developed in the text below.

Table 142

			Presupposed	Negation	Perspective	Element	Place
C _i PRES	HERE	←[PROP ...referent...]	Element	No	Speech-Act	[+SPEC,+DEF]	[−individuated]
C _i DEICTIC	HERE	←referent	Element	No	Speech-Act	[+SPEC,+DEF]	[−individuated]
C _i EXI	THERE	←[PROP ...referent...]	Location	Yes	Discourse	[±SPEC,−DEF]	[−individuated]
C _i REF	THERE	←referent	Location	Yes	Discourse	[±SPEC,±DEF]	[+individuated]

Locative sentences display (c)overt subjects (184) with topic-comment structure, where subjects raise to S_H. Raised [−SPEC] subjects make bad topics (Beaver *et al.* 2006; Bentley 2010), hence *un gatto* is questionable without context.

*Ci*_{LOC} displays narrow (argument-)focus with primary pitch accent on post-verbal subjects (185). When DPs raise, becoming the topic, [*ci*+copula] takes focus (186). If present, PP_{LOC} must be prosodically and syntactically dislocated (Leonetti 2005:10), with *ci* acting as a resumptive clitic. In questions, *wh*-phrases take clausal focus and cannot be doubled by *ci* if locative (187). The same sentences are acceptable with indefinite (194) or non-referential [−SPEC] DPs (195) with existential readings. When *ci* is referential, DPs may be [±SPEC, ±DEF]. *Ci*_{LOC} is a referential anaphor representing discourse-salient locations. Its point of origin is discourse-*here* (not speech-act), hence the reference is always distal in nature. Without discourse-salient location, *ci*_{DEICTIC} becomes the here-and-now of the speech-act with deictic reading which requires [+SPEC] DPs (188/189), since it is logically impossible to point out [−SPEC] objects. If PP_{LOC} is not dislocated, locative readings are unavailable, and interpretation is determined as existential/presentative.

Table 143

		S _H	P	N	O	V		
184		[G./ ² un gatto] _i _{TOP}		Ø _i	Ø _j	è	in giardino _i /qui _i /lì _j	[PRED-FOC] G./he is (t)here (, in the...)
185	L _j	Ø _i		Ø _i	c' _j	è	[G./un gatto] _i _{FOC} (, in giardino)	There is G. (,in the...)
186		[G./ ² un gatto] _i _{TOP}		Ø _i	[c' _j	è] _{FOC}	e _i	G. is there
187	Dove _j	Ø _i		Ø _i	*ci _j	sei	tu _i ?	Where are you?
188	Ø _j	Ø _i		Ø _i	c' _j	è	[G./ ² un gatto] _i _{FOC}	(T)here is G.
189		[G./ ² un gatto] _i _{TOP}			[c' _j	è] _{FOC}	e _i	G. is (t)here
190	L _j	<molte ragazze _i >	Ø _{non}	Ø _i	ci _j	sono	<molte ragazze _i >	Many girls are (not) (t)here
192	Ø _j	Ø _j	Ø _{non}	ci _j	Ø	sono	[molte ragazze _i]	There are...many girls
193							... <u>few</u> girls	... <u>few</u> girls
194	Dove _j	Ø _i		ci _i	Ø _j	siano	molte ragazze?	Where are there many girls?
195						è	il telefono?	...is there a telephone?
196				c' _j		è	[un gatto] _{FOC} [in giardino]	There is a cat (in the...)
197				[c' _j		è] _{TOP}	[un gatto (in giardino)] _{COMMENT}	
198	[L _j] _{NEW} ,			c' _j		è	[un gatto] _{FOC} [e _j]	In the garden _{NEW} , there is...
199								*...(t)here's the/a cat
200		Ø _j		ci _j	ne _i	sono	[molte e _i]	There are many of them
201	[Pane _i] _{TOP}			[ce	n'	è	(poco) sul tavolo] _{FOC}	There's (a little) bread on the table

Ci_{EXI} is not only compatible with PP_{LOC} , but presupposes locations, as stage topic upon which the existence of its indefinite DP is predicated (Partee & Borschev 2002, 2007; Koontz-Garboden 2009). Without PP_{LOC} , (196) is read with [–SPEC] location (\approx existence) with *ad sensum* intersection with discourse-*here*, in that the DP's existence is presumed relevant to discourse. Ci_{PRES} references the here-and-now of the speech act, and hence the objects pointed to must be [+SPEC]. In both cases, additional locatives restrict the relevant value of (T)HERE. Such locations do not co-index ci , need not be dislocated (196), and can be extracted without the need to resume them in the main clause (198).

Unlike English, Italian presupposed/old information must be dislocated (Cruschina 2012). Aboutness (often new) topics are fronted, while familiarity/referential topics may be left- or

right-dislocated. When PP_{LOC} constitutes an aboutness topic in existentials, it appears sentence initially as clausal topic (198), with no pragmatic/semantic affect. Thus, PP_{LOC} is *not* part of the focus, and existentials cannot be subsumed under presentationals where the whole sentence is ‘presented’ as new (contra Lambrecht 1994). Neither deictic nor presentative readings (199) are available since *ci*=here-and-now is what the sentence is about; it does not reference prior locations/topics.

Ci-locatives may express topic-comment variation (place-entity vs. entity-place) by raising its subject. Scope indicates different structures (Leonetti 2005:7). In locative readings (191), negation scopes over the predicate ‘*aren’t there*’ not the DP, regardless of its position. In existentials (193), negation scopes over *many*, creating *few girls*. Context determines the reading of surface-identical (190-191)~(192-193). DPs are predicates in existentials, but subjects in locative predications (197). In existentials, the DP may be extracted to pre-clausal topic position (201) and its class *ne*-extracted to OBL (200), but the DP never raises to S_H.¹¹⁶

Pragmatically, presentative constructions introduce new propositions, the whole clause taking sentence-focus. DPs are post-verbal functioning as topics of adjectival predicates (202) or pseudo-relative clauses (203), often introducing surprising events demanding focus (204).¹¹⁷ Presentatives are independent of discourse, carrying no presuppositions allowing them to be used in out-of-the-blue contexts, and preceded by questions (*What happened?*) or exclamations (*Guess what!*), which require sentence-focus replies (Lambrecht 1994:164).

¹¹⁶ We take categorial constructions e.g. Italian *Dio c’è*, ‘God exists’, with focused existential predicate as pre-verbal topics.

¹¹⁷ ‘Eventitives’ have many definitions. Berruto (1986:67) restricts the term to cases where events are expressed by single DPs (204) and predicates are equivalent to ‘happen’.

Table 144

202	C'è [_{SC} Gianni infuriato/nei guai]	John is furious/in trouble
203	C'è [_{SC} un signore [_{CP} che vuole parlare con te]]	A gentleman wants to talk to you
204	C'è [_{SC} il terremoto]	An earthquake is happening
205	Anche G (*c') è infuriato	John too is furious
206	<Anche G> c'è <Gianni> in giardino	John too is in the garden
207	[CA] Hi ha la Maria {molt enfadada/al telefon/que espera}	M. is {very angry/on the phone/waiting}
208	C'è [il Signor P che chiede di essere ricevuto]	[NEW] Mr P. asks to be received
209	C'è il Signor P [che chiede di essere ricevuto]	Mr P. here asks...
		Mr P. is here, asking...
		Here is Mr P., who...
210	C'è il Signor P, in salotto, [che chiede di essere ricevuto]	Mr P. is (there) in the living room, asking...
211	Ci fu una disgrazia	There was an (unfortunate) accident/ An (unfortunate) accident occurred

Although (203) might be translated as ‘*There is a man here who...*’, presentatives are not existentials. Although both introduce new referents, existentials introduce elements with/out predicate whilst presupposing locations, but presentatives introduce whole predications without requiring locative anchoring (202). Existentials are limited to stage-level adjectival predicates (Milsark 1974; McNally 1992), whilst presentatives also allow individual-level predicates (202). Finally, existential DPs must be indefinite, but are unrestricted in presentational sentences. Equally, presentatives ≠ locatives. Unlike locatives (206), presentatives sentence-focus cannot be broken. When DPs take argument-focus through focus-fronting, e.g. adding *anche* which requires DP narrow-focus under a contrastive/surprise interpretation (Cruschina 2012), *ci* must be omitted (205). In languages exhibiting auxiliary-change (207), presentatives take HAVE, unlike locatives.

Many cases are ambiguous (Berruto 1986:71). Presentative (208) introduces the whole proposition, ‘it is that [...]’, where Mr P may/not be present (e.g. in an anteroom), but must be ‘imminent’ to here-and-now. It may also take a locative reading; deictic (209) or referential (210) with right-dislocated location. Whilst (168) can only be presentative due to DP_[+DEF],

(211) may be read either way. Particular properties restrict possible interpretations, but selection from remaining readings must be made within discourse and speech-act contexts.

Ci_{PRES} does not reference external objects or predication settings. It is discourse-internal, *pointing out* new propositions as pertinent to the current setting; a function characteristic of narrative/spoken language (Berruto 1986). $Ci_{DEICTIC}$ *points out* objects in the current setting. Both impose speech-act HERE. Additional locations further specify the object's position *within* HERE. Introduced elements cannot be aboutness topics nor take prominence by pre-verbal topicalization, since this would clash with HERE (*ci*) which is what the constructions are 'about'. Both presuppose the introduced element and, therefore, cannot be negated (212-213). Such 'tangible', elements must be [+SPEC].

Table 145

212	*Non c_{PRES} è [Gianni infuriato/nei guai/che studia medicina]	John is not angry/in trouble/studying medicine
213	Non $ci_{DEICTIC}$ è Gianni #in giardino	#Here is not John in the garden
214	Non ci_{EXI} sono orsi bianchi al Polo Sud	There are no polar bears in the South Pole
215	Gianni, non ci_{REF} è	Gianni isn't there

Conversely, negation is acceptable with ci_{EXI} and ci_{REF} (214-215) which presuppose locations, but not necessarily the element introduced (Partee & Borschev 2007) which, therefore, may be [\pm SPEC]. Both reference THERE discourse here-and-now. Ci_{EXI} introduces indefinite objects/classes as existing in ontological space, potentially refined by additional locations. Ci_{REF} references salient places from discourse with no limitation on DP definiteness.

As demonstrated, there is a need for the four types contrasted in Table 142. Not only, must the [\pm individuated] nature of the referent be taken into account, but its relationship to the clause i.e. its case function.

5.4.4 Romanian

Romanian (examples from Cornilescu 2009) is a BE_AT language retaining dative case, but without existential/locative clitics. Existential sentences are expressed through stress/focus and display similar definiteness effects to English/Italian. Whilst (216) is a simple copular sentence with (c)overt subject, (217)'s verb is prosodically marked showing that it is (part of) the focus, i.e. 'being' is at stake, and takes existential readings. Verbal focus may be indicated through intonation (217), and/or negation (220) or focusing particle (222, *mai*). The DP may be extracted to TOPIC position (219/221), separated from the verb by a pause, leaving only BE in focus. This position is not S_H as shown by the fact that it cannot be discourse initial, but must continue a discourse where the proposition is denied/questioned (218/219). Focus indicates presence of a Ø_{EXI} subject.

Table 146

216	[Ei/Aceștia/Ø Ø _{NOM} sunt mari compozitori]	They/These are great composers
217	[Ø _{EXI} SUNT mari compozitori]	There are great composers
218	Muzica simfonică se află în declin, deși...	Symphonic music is declining, although...
219	...Mari compozitori _i [Ø _{EXI} SUNT e _i]	...great composers, there are
220	Nu este dreptate	There is no justice
221	Dreptate#nu este	Justice, there is not
222	Mai este onestitate	There still is honesty

Unlike vP-internal NPs, post-verbal NPs in existentials must take narrow scope with respect to clause-level operators e.g. negation. (223)'s post-verbal DP is an argument (subject) scoping above or below negation. In existential sentences (224), post-verbal NPs are understood only within the negated predicate; thus, (224) predicates the property of [*not many*] about *problems*. The subject is Ø_{EXI}. This abstract location may be constrained by adverbial locatives or speech-act deictic features. In (225), the indefinite space intersects with

Romania; in (226), discourse-*here* is implicit. When locatives are present or implicit, there is no focal stress on the verb (225/226). When more general spaces are envisaged (making an ontological claim) verbal stress (227) or other indicator is required. Unlike existential sentences, locative sentences are unconstrained regarding possible subjects and position. They may be initial in discourse, [\pm definite], unfocused, and require no stress (228/229).

Table 147

223	N-au venit mulți studenți	Not many students came/Many students didn't come
224	Ø _i nu sunt multe probleme	There _i are [not many problems]
225	E secetă în România	There is draught in Romania
226	E secetă	There is-draught (here)
227	ESTE foamete	There is hunger=there are places afflicted by hunger
228	[Studenții Mariei] _i Ø _i Ø _i sunt [în clasă] _i	Mary's students are in the classroom
229	[{Unii/Cealalți} copiii] _i Ø _i Ø _i sunt [la cinema] _i	{Some/the other} children are at the cinema

Thus Romanian has the same range of clitics/functions as other Romance languages; they are merely silent. Their presence is evidenced by structure, and where necessary intonation.

5.4.5 Sardinian

Sardinian¹¹⁸ highlights existential vs. locative clitics by change in copula (Jones 1993; La Fauci & Loporcaro 1997; Loporcaro 1998; Bentley 2004, 2011; Remberger 2009; *i.a.*). In locative sentences, definite DPs (including (c)overt personal pronouns, 232) select BE_{AUX} with verbal agreement (230-234). As indicated by personal pronouns, definite DPs take nominative, appearing pre-/post-verbally (230, 232). *Bi* and PP_{LOC} are mutually exclusive (233). To appear in the same sentence, PP_{LOC} must be dislocated (234), as indicated by intonation, but not always orthographically (Remberger 2009). *Bi*'s referent must be recoverable.

¹¹⁸ Examples from Jones (1993:100, 3.2.2, 113, 3.2.4)

Table 148

	S _H	N	O	D	A	Aux		
230	<Zubanne _i >	Ø _i	(b')			est	arribatu <Zubanne _i >	John arrived (there)
231		Ø _i	(b')			sun	sas pitzinnas _i	The girls are there
232	<Nois _i >	Ø _i	(b')			semus	<nois _i >	We are there
233	Ø _i	Ø _i	<bi> _j			soe	arribatu <a domo> _j	I arrived {there/at home}
234	Ø _i	Ø _i	bi _j			soe	arribatu, [a domo] _j	I arrived there, at home
235							[tres <pitzinnas _i >]	There are 3 {girls/of them}
236	Ø _j	b _j '			<in _i >	at	arribatu [tres <pitzinnas _i >]	There arrived 3 {girls/of them}
237							ballatu [tres <pitzinnas _i >]	There danced 3 {girls/of them}
238		bi _j	nke _k		nd' _i	at	issitu [tres Ø _i]	There came some three _i out of there _k

Existential sentences present the existence of indefinite objects (235), or unaccusative (236) and unergative (237) events. *Bi* is obligatory even for weak existential readings, as is HAVE_{AUX} which ‘agrees’ with its null-subject i.e. default 3.SG. The DP must be indefinite (allowing *ne*-extraction) and post-verbal;¹¹⁹ it cannot raise, since S_H is already filled. It follows that inherently definite personal verbal forms cannot appear in event-introducing existentials. *Bi* has no referent other than ontological space, but *ne*_{OBL} must always be [+referential] and may co-exist with *ne*_{PRT} (238).

Table 149

	S _H	N	O		Subject
239	Ø _i	Ø _i	ci _j	sono arrivati (, a Roma _i)	They _i arrived {there/at Rome}
	Ø _i	Ø _i	Ø _j	sono arrivati (a Roma _i) tre uomini	Three men _i arrived (at Rome)
					There arrived three men _i (at Rome)
240					
Proximal		Sardinian	Italian	French	Romanian
		bi	ci	y	Ø
Medial		bi	vi	y	Ø
Distal		bi	Ø	Ø	Ø

Whilst Sardinian has surface *bi* for all constructions, French/Italian do not use *ci/y* in presentationals i.e. weak existentials (239, Leonetti 2005, 8). We propose that this derives from lexical differences in each language’s proximal~medial~distal clitic lexicon (240), yet

¹¹⁹ When pre-verbal, it is no longer a thetic construction but categorial as indicated by auxiliary: *Tres pitzinnas (bi) sun vennitas* (Jones 1993:102). This is similar to Italian (fn.116, p.247).

again underscoring the work's central tenet of focusing on function, *not* form.

5.4.6 Diversity of *Ci_{EXI}*

Many assume that the DP is the main predicate of existentials (cf. Williams 1994; Hazout 2004; Francez 2007) whilst its topic/subject is a location (cf. Babby 1980; Partee & Borschev 2002, 2007; Leonetti 2008). Independently of the presence of PP_{LOC}, the argument of the property denoted by existential DPs is always an implicit contextual domain (intuitively similar to location), where overt locative codas contribute to the restriction of its identity (Francez 2007). The concept of a null/implicit location as the argument of existential predications has been formulated in various terms e.g. stage topic (Erteschik-Shir 1997), event argument (Kratzer 1995), and identified with null locative arguments postulated for unaccusative constructions (Benincà 1988; Saccon 1993; Pinto 1997; Tortora 1997, 2001; Sheehan 2006, 2010). The pro-form has been considered an impersonal/expletive subject (Spanish, Suñer 1982; Catalan, Rigau 1997, 1994), a quasi-argument as in weather expressions (French, Kayne 2008), and as arbitrary *pro* with non-referential reading (Cabredo Hofherr 2006).

Williams (1994), Hazout (2004), and Francez (2007) *i.a.* analyse existential DPs as predicate nominals; *there/ci* is an 'expletive' subject, originating in subject position of existential small clauses, raised to S_H. Similarly Bowers (1993) and Remberger (2009), using predicative phrase structures. *Ci_{EXI}*, however, appears within the clitic-field (following *non*), not in S_H, and therefore, cannot be an 'expletive' subject in these terms, as confirmed by French which combines expletive subject *il* (S_H) with Ø_{NOM} and *y* (OBL). Similarly, in Old Tuscan, overt expletive *egli* accompanies *ci_{EXI}* (Ciconte 2010). Mensching & Remberger (2006) for other Romance varieties. Subject (S_H) must, therefore be separated from NOM clitic (241).

Table 150

241	S _H	N	S _H	N	S _H	N	S _H	N	S _H	N	S _H	N	
Presentative	Ø	Ø	Ø	Ø	Ø	ci _{PRES}	Il	Ø	Ø	Ø	Ø	bi	[+SPEC]
Existential	egli	ci _{EXI}	Ø	ci _{EXI}	Ø	ci _{EXI}	Il	Ø	Ø	Ø	Ø	bi	[-SPEC]
	Old Italian		Italian		Spoken Italian		French		Romanian		Sardinian		

242	S	N	O	French		S	N	O	Italian		S	N	O	Sardinian		SPEC
Personal	Il			pleure					piange					pranghende		+
Locative	Il	<y _j >		est arrivé <à la maison _j >					<ci _j >sono arrivati <a casa _j >					<b _j '>soe arribatu <a domo _j >		±
Expletive	Il			pleut					piove					est proende		–
Existential	Il	y	Ø _j	a 3 hommes (à la maison _j)			ci	Ø _j	sono 3 uomini (a casa _j)			b'	Ø _j	at 3 òmines (a domo _j)		–
Weak Exi.	Il			est arrivé 3 hommes					sono arrivati 3 uomini			b'		at arribatu 3 òmines		–
Presentative							ci	sono arrivati gli uomini								+

One approach (in line with *ne*'s analysis, §5.3.1) sees the DP as the clausal subject in all cases. Merged as *v*P's external argument, it checks its features in TP/IP (causing verbal agreement) including setting SPEC,IP as [\pm SPEC], but does not raise to S_H if it is non-agentive/[–SPEC], as in existentials/presentatives. S_H and NOM are spelt-out using language-specific lexical entries for each feature combination, often Ø. Since most languages have not developed *ci*_{PRES}, the resulting spell-out with Ø results in sentences interpreted as locative constructions where possible, or ungrammatical, where not. In existentials/presentatives, additional locations map onto OBL but cannot pronominalize since this would create a sequence of *ci*'s, breaking RND. This approach is able to represent all constructions in all language types (242); for Romanian everything is Ø.

However the featural details of these functions are formulated, it is clear that, in addition to *ci*_{DAT}, there is a need for *ci*_{NOM} and *ci*_{OBL}, each able to reference real or abstract entities. Unlike *ne* with its four cases, the ACC form mapping to this category is represented by particular uses of *lo/la* (§5.5.1).

5.4.7 Exclusions

Despite the fact that not all combinations appear on the surface, we argue that there is no clitic~clitic exclusion mechanism.

Since individuals cannot be affected by the mere existence of a class of objects, only by a specific set of them, personal OBL (i.e. individual event affectees) are not available in existentially interpreted sentences. This is a matter of logical interpretation. Similarly, since ne_{ABL} references particularities rather than existential classes, only ne_{GEN} , which selects a part of that class, is available. Thus, $ci_{EXI}+\emptyset$ and $ci_{EXI}+ne_{GEN}$ are the only logically possible surface combinations.

Subject-oriented ci_{OBL} is inappropriate with SE_{NOM} which highlights the predicate's object-oriented perspective; $SE_{NOM}+ci_{DAT}$ is fine. SE_{ANT} defines initiation of new states, with optional reference to source (ne) or affectee (OBL). Such constructions specifically denote COS making ci_{OBL} 's stasis inappropriate, hence $*SE_{ANT}+ci_{OBL}$.¹²⁰ $SE_{ANT}+ME$ etc. are not possible in Italian/French due to an independent language-specific limitation (§3.3.5), but common in Spanish (§3.3.2). Similar arguments hold for $*ne_{NOM}+OBL$, $*ne_{NOM}+ne_{ABL}$, and $*ne_{NOM}+ci_{OBL}$, whilst we assume that $*ne_{NOM}+ne_{OBL}$ is a 3-3-restriction as found in the lower clitic-field. For verbs describing changes of disposition rather than position, $SE_{ANT}+ci_{OBL}$ should be available. Whilst it is found in languages with y/i forms (e.g. Aragonese, where it is so common as to be described as 'pleonastic', §6.6), we found no examples in Italian corpora, although Pescarini (2015) considers that this can occur (§3.3.3).

¹²⁰ As discussed in §5.5.4, even remaining in a state is measured by ne .

All these cases might be viewed in terms of semantic features limiting available syntactic structures, but this would be external to clitic syntax/morphology. Many of these restrictions can be derived from $[\pm E]$ e.g. SE_{ANT} ($[-E]$) only appears with $OBL_{[+E]}$, SE_{NOM} ($[+E]$) with $OBL_{[-E]}$. Further investigation (particularly with regard to the upper clitic-field and those languages which support non-reflexive nominative clitics) would be required to show if this held across all situations. Even if this proved to be the case, however, it would not indicate a clitic~clitic restriction, but merely reflect the existing semantic restrictions which allowed the construal to be formed and later presented in syntax. Absence of these combinations merely reflects higher levels of language. There is no evidence for ‘feature arithmetic’ or clitic~clitic restrictions other than 3-3-contexts, which are discussed in Chapter 6.

5.4.8 Conclusions

The conclusions are very simple. There are more clitics, each with more specific uses and hence positions, than most theories cater for; summarised in Table 151. Once this is accepted, there are no combinatorial restrictions to account for. Moreover, as will be shown below, there is no difficulty in compositionally interpreting them.

Table 151

243	NOM		OBL		DAT		ACC	
	\emptyset	NE_{PRT}	NE_{GEN}	NE_{ABL}	NE_{GEN}	NE_{ABL}	NE_{PRT}	
	$[-R]$ $[\pm I]$	CI_{EXI}	CI_{IMP}	CI_{LOC}	CI_{IMP}	CI_{LOC}	$LO_{PHRAS/LA_{ABS}}$	
	$[+R]$ $[+I]$	SE_{ANT}	SE_{NOM}		SE_{MID}	SE_{DAT}	SE_{PASS}	SE_{ACC}
	$-E$	$+E$	$-E$	$+E$	$-E$	$+E$	$-E$	$+E$

5.5 Putative ‘Lexicalization’

§5.1.1 showed that many clitic uses are (randomly, from our perspective) selected for lexical storage. Below, the analysis of previous sections is applied to such cases, showing how this is unnecessary, indeed, misleading.

5.5.1 $Lo_{PHRASAL}/La_{ABSTRACT}$

English has numerous expressions containing *it* which may reference a range of propositional types. Without neuter forms, Italian must express *it* as a masculine or feminine clitic. Rather than see verbs such as *capirla* and *cavarsela* as special cases stored separately in the lexicon, they should be seen as simple transitives with an *it* for their direct-object; they are no more ‘pronominal’ than English ‘get it’ = ‘comprehend’.

Following on from §5.1.2, *lo/la* can be [+individuated] lo_{ACC}/la_{ACC} or [–individuated] $lo_{PHRASAL}/la_{ABS}$. Whilst *lo* may be used to anaphorically reference clausal propositions (246-247, Maiden & Robustelli 2000), *la* expresses (244-245) abstractions pertinent to the context. *La* is often seen as referring to ‘covert’ feminine NPs, recoverable from context or inherent in the verb’s semantics e.g. *una storia (raccontarla, 248)*, *una situazione (prenderla, 249)*, or generalized objects (*una/la cosa*) often used to avoid taboo e.g. *farla* ‘defecate’, *darla* ‘of a woman, have sex easily lit. gives it’. In other cases, however, it derives from Latin N.PL ILLA (Rohlf 1968:§456) used to express collections¹²¹ e.g. things in *pensarla* (250-251).¹²² Whilst both *lo* and *la* may be considered ‘neuter’, *lo* is [SG,+DEF] (with the clause as specific referent), whilst *la* is [–SPEC]. It is inconsistent that la_{ABS} should be used as evidence for lexicalization, but not $lo_{PHRASAL}$.

121 Ancient collective number, as expressed by N.PL subjects, took singular verbs (Sihler 1995). Vestiges of this arrangement remain. Italian plural forms distinguish between *ossi* (bones, conceived separately) and *ossa* (set of bones/skeleton) corresponding to the collective meaning (Spitzer 1941:341). Romanian possesses a category of (surface feminine) nouns with abstract denotation, “whose plurals have collective meanings or refer to different types of the objects designated” (Hall 1965:424).

122 For use as ‘potential resolution’ of the current SOA, §5.5.3.

Table 152

244	Piantala!	Cut it out! [lit. Plant it!] e.g. that behaviour
245	Non la capisco!	I don't get it! e.g. the answer to a problem
246	[Oggi è festa] _i , non lo _i sapevi?	[Today's a holiday] _i , didn't you know [it/that] _i ?
247	[La pianura era spesso avvolta nella nebbia] _i , ma quel giorno per fortuna non lo _i era	[The plain was often shrouded in fog] _i , but luckily that day it wasn't [so] _i
248	A chi la racconti?	Who are you trying to fool?
249	I tifosi la prendono bene/male	The fans take it well/badly
250	Ha scelto me per come la penso, non perché...	They chose me for my opinions, not because...
251	Una società totalitaria era perseguitato chi la pensava diversamente	A totalitarian society persecutes the people who think differently

Whilst some noun-replacement readings may be historically accurate (e.g. *battersela* < *battere la ritirata*), (non-)inclusion of *la* is often pragmatically driven and, therefore, not a matter of necessary syntactic realization of objects, but rather a choice between two construals. According to Russi, native speakers find *finirla/smetterla* to be “stronger”; expressing the speaker's emotional involvement, e.g. speaker irritation with unresponsive addressees (252). Without *la*, utterances lose their unpleasantness, and may become pragmatically inappropriate. In (253), *la* expresses speaker affectedness due to the addressee's actions, whilst in (254), *la* would be unusual for someone expected to maintain professional distance. Conversely, *la* is impossible in (255, taken from recipes), because the speaker/narrator cannot possibly be affected by the event.

Table 153

252	Smettila di scusarti -proruppe lei. -È accaduto e basta	Stop apologizing, she burst, It happened and that's it
253	Io le sono affezionata. Lei dovrebbe smetterla di bere	I care about you. You should quit drinking
254	«Devi smettere di bere» lo aveva ammonito il medico	'You must quit drinking' the doctor had warned him
255	Aspettate 2 minuti, finché le patate smettono di emettere vapore	Wait for 2 minutes until the potatoes stop steaming
256	Quando la cominci con queste scemenze ti prenderai a schiaffi	When you start (it) with this foolishness, you make me want to slap you

At a semantic/pragmatic level, *la* increases subjectivity representing the speaker's perspective in discourse (“speaker's imprint”, Finegan 1995:1). Syntactically, however, *la* (when present)

is simply an expression of the accusative argument; *la attività di+infinitivo* ‘the (activity) of...’, which is present even when covert. The difference is analogous to English *Stop whining!*~*Stop it with all this whining!* Association with particular verbs is register-based; e.g. *la* is not found with *terminare/cessare* which are less frequent and largely restricted to higher, more specialized registers, but may be used (less systematically) with *cominciare* (256). The more formal situations which require these verbs also militate against the use of personal indicators. The speaker, therefore, has a choice between \emptyset_{ACC} and *la_{\text{ACC}}*. It is not determined by lexical entries.

5.5.2 *Se+Lo/La*

Transitive hosts of accusative *lo_{\text{PHRAS}}/la_{\text{ABS}}* may take further arguments in order to compose a desired meaning. Thus, *immaginare* ‘to picture’→*immaginar+si_{\text{DAT}}* ‘to picture for oneself, imagine’, whilst imagined objects may be real or previous propositions (257). Equally verbs taking *la_{\text{ABS}}*, may also take personal (258) or adverbial (259) clitics. If the dative happens to be 3.REFL, *~sela* is formed (260); just like *~cela* (259), or *~selo* (257). This is simple composition, requiring no special treatment.

Table 154

257	Non riusciva di immaginarselo	She couldn’t even imagine it	Proposition
258	Tu non me la dai a bere	You don’t fool me (<Give it to me to drink)	Abstraction
259	Non <u>ce la</u> racconti giusta	You are not telling the truth about that	(cf. 248)
260	<u>Se_{\text{DAT}}</u> la prende per niente	He takes offence for nothing	(cf. 249)
261	<u>Me_{\text{NOM}}</u> la prendo con te	I take it out on you	

Cavare requires direct-object and source complement (262), which may be implicit (263), or recoverable from context. *Cavarsi* can be construed as direct-reflexive (‘free oneself from a difficult situation’, 264), or as indirect-reflexive where the locative source is the subject’s

personal domain i.e. *si* is possessor of the direct-object, whether concrete (265) or abstract (266). *Cavarsela* is compositionally ‘pull it off for oneself’, where ‘it’ (*la_{ABS}*) is the pronominalized direct-object referencing a successful conclusion (from the subject’s perspective, hence *si*), to the current SOA of which the subject takes possession (267); often translated ‘manage’ (268).

Table 155

262	Ho cavato [dalla tasca] _{LOC} [il portafogli] _{DO}	I pulled out my wallet from my pocket
263	Intanto dovevo farmi cavare [il dente] _{DO} ,...	Meanwhile, I had to have my tooth pulled out
264	Ha pensato a cavarsi dai guai	He took care to get out of trouble
265	Il vecchio mugnaio si cavò rispettosamente [il berretto]	The old man took off his hat, respectfully
266	Si è cavato [il capriccio] _{DO} di comprarsi una Ferrari	He was satisfied his whim of buying a Ferrari
267	Se ce la caviamo,...	If we manage/get out of this/pull it off ...
268	Me la cavo più o meno in tutte le materie	I manage more or less in all subjects

Some verbs produce *~sela* by applying *si_{NOM}* (261), the subject involvement of which matches *la_{ABS}*’s subjectivity (§5.5.1), whilst adding notions of energy and completion/satisfaction, not found with patient-oriented *si_{DAT}*, where the reflexive references *la_{ABS}*’s affected possessor within an unfolding state (260). For *Aver((se)(la))* see (350-351, p.270). Some verbs show all uses (269-272) including literal readings when a clear anaphor is present (273).

Table 156

269	Battere+Ø la ritirata	Beat the retreat	Signal exit for others
270	Batter+se _{DAT} la ritirata	Beat the retreat for oneself	Exit under own compulsion
271	Batter+se _{NOM} la ritirata	Beat a hasty retreat	SE _{NOM} => energy/completion
272	Batter+se+la	Beat it (hastily)	LA _{ABS} subjectivity
273	La porta, se la batte furiosamente	The door, he beat it furiously	Anaphoric reference

Constructions available to a verb (*Ø~la~si~sela*) are defined by verbal semantics; whilst appropriateness is determined by context. These are not special cases which require lexical storage; the meanings remain compositional. There is no *prendersela*, *battersela*, etc., just as there is no *immaginarselo* or *raccontarcela*. Isolating such uses is unjustified. All that is necessary, is to recognise their components.

5.5.3 Object-Oriented *Ce+La*

When no discourse-salient location is present, *ci* defaults to readings of discourse-*here*, not only in the sense of a physical place but also as current SOA or proposition, where union of object and state has the potential to change that state, leading to a new discourse-*here*. This is frequently combined with *la*_{ABS} representing the ‘resolution’ being brought to, lacking from, or possessed at, that situation.

Metterci represents the locating of concrete physical entities in the spatial domain (274-275), including oneself (276), or application of an abstract entity (often represented by *la*_{ABS} (277)) into the current SOA (*ci*), construed as an abstract place (278). The most common abstract objects are time expressions (279-280). If present, *metterci*’s second object is clausal (279) with coreferential subject. *Ci* does not substitute/double this clause but represents the current SOA as a place where putting the abstract object will lead to that clause’s realization. In (279), focus is upon subject injection of effort into the situation, whereas the *showering* is almost incidental. In (280), the outcome is not even mentioned but inferred from context. When an [+individuated] place is present, *ci* must be read as resuming it (275), otherwise it defaults to discourse-*here*.

Table 157

274	Carlo mette le chiavi nel cassetto	Carlo puts the keys in the drawer							
275	[Nel cassetto] _i , Carlo ce _i le mette	Carlo puts them there in the drawer							
276	Mettersi in movimento	=è partito come un fulmine							
277	Mettercela tutta	To put everything into it/give it one's all							
278	Ci devi mettere piu energia	You must put more energy into it							
279	Carlo ci mette dieci minuti [a farsi la doccia]	Carlo takes ten minutes to shower							
280	Ci hai messo una vita!	It took you ages! (lit. "a lifetime")							
	<table><tr><td>X_{SUBJ}</td><td>{</td><td><table><tr><td>puts Y_{OBJ}</td></tr><tr><td>puts <i>it</i> (=la)</td></tr></table></td><td>}</td><td>in W_{situation} (=ci) [so that Z_{CLAUSE} is/becomes true]</td></tr></table>	X _{SUBJ}	{	<table><tr><td>puts Y_{OBJ}</td></tr><tr><td>puts <i>it</i> (=la)</td></tr></table>	puts Y _{OBJ}	puts <i>it</i> (=la)	}	in W _{situation} (=ci) [so that Z _{CLAUSE} is/becomes true]	
X _{SUBJ}	{	<table><tr><td>puts Y_{OBJ}</td></tr><tr><td>puts <i>it</i> (=la)</td></tr></table>	puts Y _{OBJ}	puts <i>it</i> (=la)	}	in W _{situation} (=ci) [so that Z _{CLAUSE} is/becomes true]			
puts Y _{OBJ}									
puts <i>it</i> (=la)									

Far((ce)(la)) follows a similar pattern. *Ci* in (281) is resumptive. In (282), it references an SOA (discourse-*here*) perceived as ‘in need of resolution’. *Ci* is not obligatory, but omission weakens this inference (283), as indicated in the translations. Adding *la*_{ABS} (284/285) creates readings of ‘manage/succeed’, where *la*_{ABS} refers to the SOA’s *resolution* i.e. ‘whatever is necessary’ as defined by context. Neither clitic references the optional *a*+INFINITIVE clause selected by *fare* (285), i.e. the desired SOA₂ following SOA₁’s resolution. Whilst *metterci* highlights *what* is being put into the situation, *farci* highlights the action itself.

Table 158

281	In questa situazione _i , non possiamo farci _i niente	In this situation, we cannot do anything								
282	Che poteva farci, povero Berto	What could he do about it, poor Berto								
283	So cosa far(cì) _i	I know what to do (about it) _i								
284	Il pilota è formidabile. La Ferrari potrà farcela...	The pilot is exceptional. Ferrari can make it...								
285	Non ce la faccio [ad essere sempre il più bravo]	I can't manage to be the best all the time								
	<table><tr><td></td><td>X_{SUBJ}</td><td>{</td><td><table><tr><td>acts</td></tr><tr><td>does it (=la)</td></tr></table></td><td>}</td><td></td></tr></table>		X _{SUBJ}	{	<table><tr><td>acts</td></tr><tr><td>does it (=la)</td></tr></table>	acts	does it (=la)	}		in W _{situation} (=ci) [so that Z _{CLAUSE} is/becomes true]
	X _{SUBJ}	{	<table><tr><td>acts</td></tr><tr><td>does it (=la)</td></tr></table>	acts	does it (=la)	}				
acts										
does it (=la)										

Avere functions as auxiliary and main possessive verb. *Averci* is widespread (286), considered part of *italiano popolare* (Battaglia & Pernicone 1968:154), or colloquial Italian (Sabatini 1985:160). D’Achille (1990) provides examples from the XIV^c where *ci* retains referential value, and of ‘true’ *averci* from the XVI^c. Pulgram (1978) foresees lexical divergence whereby *avere* will survive as auxiliary only, and *averci* become the verb of possession; cf. Spanish *haber* (auxiliary) vs. *tener* (possession). Many Italo-Romance varieties have similar constructions (La Fauci & Loporcaro 1993, 1997; Moro 1998; Benincà 2007).

Table 159

286	C' ho un formicolio alle mani	I have a tingling in my hands						
287	(C') hai le chiavi? –No, non [?] (ce) le ho	Do you have the keys? –No, I don't have them						
288	Ma ha ragione ad avercela con i giornalisti	But he is right to be angry at journalists						
289	Ho avuto i primi sospetti che qualcuno ce l'avesse con me	I began to suspect that somebody was mad at me						
	<table><tr><td>X_{SUBJ}</td><td>{</td><td><table><tr><td>holds Y_{OBJ}</td></tr><tr><td>holds <i>it</i> (=la)</td></tr></table></td><td>}</td></tr></table>	X _{SUBJ}	{	<table><tr><td>holds Y_{OBJ}</td></tr><tr><td>holds <i>it</i> (=la)</td></tr></table>	holds Y _{OBJ}	holds <i>it</i> (=la)	}	in W _{situation} (=ci) ([so that [_{TOP} <i>ec</i>] is/becomes true])
X _{SUBJ}	{	<table><tr><td>holds Y_{OBJ}</td></tr><tr><td>holds <i>it</i> (=la)</td></tr></table>	holds Y _{OBJ}	holds <i>it</i> (=la)	}			
holds Y _{OBJ}								
holds <i>it</i> (=la)								

Averci is rare with bare nominals e.g. *avere fame* which denote states rather than possessable objects, whilst its frequency increases when objects are also pronominalized (287). Whilst absence of *ci* in questions (287) has limited effect, omission from answers is considered ungrammatical (Dardano & Trifone 1995:243). Without *ci*, sentences denote general possession. Its presence emphasizes possession within, and potential effect upon, the current SOA. *Avercela* (288-289) may, therefore, be read as *current* possession of a covert NP likely to have effect on that SOA e.g. *la rabbia*; cf. English ‘to have had *it* with someone’.

Unlike the above verbs, *volerci*’s ‘resolution’ takes subject position. Following De Mauro (1999-2000)’s dating (pre 1375), *volerci* is the oldest *verbi procomplemetari* after *andarsene* and *esserci/vi*. The transitive~intransitive alternation *volere* ‘want’~*volerci* ‘be necessary, required; take (intransitive)’ mirrors English ‘the fence {wants/is in want of} a lick of paint’, where being ‘in need’ is expressed as ‘wanting to have’.

Transitive *volere* selects nominal complements (290), or acts as a (semi-)modal¹²³ selecting clauses (291). Being desiderative, subjects tend to be human and raise to S_H, reflecting their agentivity. *Volerci*’s subject is normally post-verbal and inanimate (292), most often temporal expressions (293). Like *metterci* and *farci*, *volerci* may simultaneously select a clausal (292-293) or PP complement (294) which may remain covert (295) when recoverable from context. *Volerci* is found in several fixed phrases which require no complement, being interpreted as the current SOA’s desired outcome: e.g. *che ci vuole?*, ‘what does it take?’, *ci vuole poco* ‘it takes little’.¹²⁴

123 WANT may be analysed as WANT+HAVE (Fodor & Lepore 1998) or WANT+BE+PP (Harley 2004).

124 Several verbs follow this pattern e.g. *ci manca poco* ‘there is not long to go’ lit. little is lacking (t)here.

Table 160

290	[Gianni] _s vuole [una tazza di brodo caldo]	G. wants a cup of hot broth
291	[Gianni] _s vuole [vederti/che io continui a studiare]	G. wants [to see you/me to continue studying]
292	[Per acquistare il farmaco] ci vuole [la ricetta medica] _s	To buy the medicine, the prescription is needed
293	Ci vogliono in media [sei mesi] _s [perché una pagina...]	On average, it takes six months for a page to...
294	[Per le labbra], ci vuole [uno stick dal filtro altissimo] _s	For lips, you need one with very high sun block
295	Ci vuole [un tovagliolo] _s (non il grande asciugamano...)	You need a napkin (not the big towel...)
296	Mi ci vogliono due euro	Two euros are lacking on me,=I need 2 euros
297	A Maria, ci sono voluti sei mesi per riprendersi	M. needed 6 months to get well

For W _{OBL}	X _{SUBJ}	{	is lacking (intransitive)	}	from W _{situation} (=ci) [for Z _{CLAUSE} to be/become true]
----------------------	-------------------	---	------------------------------	---	---

298	Ci voglio io/Ci vuoi tu/Ci vogliamo noi per...	I am/You are/We are needed here...
299	Ci voleva lui, Silvio Berlusconi in persona	It was SB in person who was needed there
300	Ci sono tre uova nel frigo	There are 3 eggs in the fridge
301	Ci vogliono tre uova [per (fare) questa torta]	3 eggs are needed [{for/to make} this cake]
302	Per ogni tipo di gioco c'era un edificio	For every type of game there was a building

Volerci is similar to *metterci*, whilst accepting wider ranges of entity and carrying a sense of necessity. The state of necessity (*ci*=discourse-*here*) is perceived as impersonal, but ‘needers’ i.e. event affectees may appear as OBL clitics (296) or phrases (297), which is impossible with stative existentials. Although not widely accepted (Russi 2006:253–57), *volerci* may occur with local-person pronouns (298–299). Being inherently definite, the latter cannot be existential pivots but nominative subjects with *ci* referencing discourse-*here*.

Contra Burzio (1986) and Salvi (2001) *i.a.*, *volerci* is *not* analogous to *esserci* introducing existence, and focusing upon the expression, of the ‘needed object’, but rather the SOA which lacks that object. *Essere*-existentials (300) take locative codas narrowing down the spatio-temporal circumstances for which the entity’s existence is predicated. It is always implicit, defaulting to discourse-*here*. WANT-constructions contain an implicit reason for, or intention behind, the desire optionally expressed as final (infinitival) codas (301). E_2 is possible *iff* E_1 (the desire) is satisfied. The two events are semantically and syntactically separate. *Essere*-existentials may support purpose phrases (302, Mereu 2011:120), but they are purely optional,

whereas *volere* necessarily ‘licenses’ the purpose clause as a ‘potential resultant’ clause. *Esser*+*ci*_{EXI} is purely stative. *Voler*+*ci*_{LOC} is not ‘there is a need for x’ (=static state), but ‘x is needed here’ (=active state) with a potential/resultative state (i.e. new *discourse-here*) if that need were met.

For each of these ‘special’ cases, *ce* may be replaced by personal pronouns or *se* producing *-sela* (§5.5.2), whilst *la* may be substituted by *ne* producing *-cene* (§5.5.4). Different components produce different meanings, requiring no ‘special’ place in the lexicon.

5.5.4 Object-Oriented *Ne*

Sapere ‘possessing knowledge/notions’ takes clausal (303) or nominal (304) complements. *Sapere di*+*NP* conveys ‘having expertise in a field’, or ‘notions/knowledge about something’ (305). As the verb’s ‘internal accusative’, such knowledge need not be overtly expressed, but may be modified by adjectives (*poco*, 306); resumed by *la* (310), *lo* (314), or *ne* with quantified objects (307); or question words (315). Lack of intonational breaks indicates that *di*-phrases are not products of right-dislocation, but subordinate to the ‘internal accusative’. It may be extracted to *ne*_{GEN} (315), except where it would generate 3-3-clashes.

Usage is purely compositional, using clitics appropriate for: understood (\emptyset _{ACC}), relevant (*la*_{ABS}, i.e. potential resolution), partial/indefinite (*ne*) information, or previous propositions (*lo*_{PHRAS}). *Ne* also has pragmatic effects. In statements, absence of *ne* is neutral (306), whilst presence indicates speaker evaluation of their own knowledge (307). In questions without *ne*, the speaker awaits an informative reply (311), whilst with *ne*, the speaker expects no answer, thereby invalidating listener knowledge/opinion (312). In (315), it helps defer responsibility.

In general, absence of *ne* reflects formality, whilst resumptive clitics indicate colloquial/informal registers (Benincà *et al.*:190). These are composed choices, not lexicalized items.

Table 161

	Topic/S _H	D	A		
303			Ø _i	so [che hanno avuto una bambina] _i	I know that they had a baby-girl
304	Gianni		Ø _i	sa [la lezione di storia] _i molto bene	G. knows the history lesson very well
305			Ø _i	sapevo [Ø _i di una sua simpatia per...]	I knew of his attraction to...
306		Ø _j	Ø _i	so [DP poco Ø _i [di C] _j]	I know/have little knowledge of C.
307		Ø _j	ne _i	so [DP poco e _i [di C] _j]	I know some small amount about C.
308		ne _j	Ø _i	so [DP poco Ø _i [e _j]]	I know little of/about it
309		Ø _j	ne _i	so [DP poco e _i [Ø _j]]	I know a little
310		Ø _j	la _i	so [DP lunga e _i]	I know a thing or two ¹²⁵
311	Che cosa _i	Ø _j	Ø _i	sai [DP e _i [di C] _j]?]	What do you know <u>about</u> C.
312	Cosa	Ø _j	ne _i	sai [DP e _i [di me] _j] per dare giudizi?	What do you know <u>of</u> me to judge?
313	Q : Dove			sono le forbicine?	Where are the nail-scissors?
314	A1: Non		(lo _i)	so [DP e _i]	I don't know (it=information requested)
315	A2: Che _i	ne _j	Ø _i	so [DP e _i [e _j]]?	Why would I know of/about it?

Intendere ‘understand’ takes accusatives including *la*_{ABS} ‘it/things’ (316) and forms SE-passives with inanimate subjects (317/318). Animate subjects mark personal (319, reflexive) or shared (320, reciprocal) possession of the understanding through dative clitics.¹²⁶ As an internal accusative, the understanding need not be expressed (322), but may be (321), where specific kinds of (contextually available) understanding are highlighted e.g. *intendersela* ‘make a deal/have a relationship with someone (typically illicit)’; where generic knowledge is at issue (323, i.e. expertise); or if the knowledge is overtly quantified (324, *poco*). In such cases, *la~ne* is used to highlight its collective~partial nature. Like *saperne*, *di*-phrases are object, not verbal, arguments. Unlike *saperne*, *di*-phrases never extract to DAT, since that position is filled by the possessor.

125 lit. I know it/things at length, cf. *quanto la fai lunga!*, ‘you go on and on!’ (in a discussion).

126 As SE_{MID}, an inchoative reading ‘come to an agreement’ is also available.

Table 162

	Topic/S _H	P	D	A			
316	Io	non		la _i	intendo	<i>e_i</i> così	I don't see things/it that way
317	Ø _i			s _i '	intende	<i>e_i</i>	Of course!←It _{EXPL} is understood
318	Ø _i			s _i '	intende	[che verrai anche tu] _i	It _{EXPL} is understood that you'll be coming too
319	Con lui		mi	Ø _i	intendo	[<i>e_i</i> benissimo [Ø]]	I have a fine understanding with him
320			ci	Ø _i	intendiamo	ottimamente	We understand each other perfectly
321			se	la _i	intende	<i>e_i</i> con M	He is having an an affair with M.
322			se	Ø _i	intende	[<i>e_i</i> Ø [di musica]]	He has a lot of knowledge (about sth)
323			se	ne _i	intende	[Ø _i [di musica]]	He has knowledge/expertise (of music)
324			se	ne _i	intende	[poco <i>e_i</i> [di X]]	I know very little (about it)

Russi accepts that *ne* is not always required, seeing this as evidence that *intender* is not yet fully lexicalized. On the contrary, we argue that *ne* is no more, or less, ‘obligatory’ for *intender* than *saper* etc.. Differences in underlying argument structure determine which variations are available. *Intender* requires *se*_{DAT} to indicate ownership of the understanding, thereby denying options which include *ne*_{GEN} under DAT (308,315), but like *saper* may express its accusative as Ø, *la*_{ABS} or *ne*_{PRT} (pragmatic/register-dependently), leading to *~sela* and *~sene*, as appropriate to the meanings being composed.

There exists a range of verbs *fregar(se(ne))*, *fotter(se(ne))*, *infischiar(se(ne))*, *sbatte(re(ne))*, with numerous regional and/or register-dependent variants with personal and impersonal constructions, broadly translated as ‘I don’t care/give a damn’ which Russi considers fully lexicalized. The most widely used/acceptable is *fregare*, ‘rub, pinch, scour’; *fregarsi*, ‘rub oneself/each other’; *fregarsene*, ‘to not care’ i.e. ‘not rub/irritate/bother oneself about’.

The personal construction (325-328) ‘requires’ SE in order to express subject involvement; without it, simple transitive readings apply. Similarly, *ne*_{ACC} is ‘required’ to reference

[–individuated] abstractions related to the verb root. Since each verb has sexual connotations, non-expression of complements follows from taboo. The degree of indifference may be quantified (334), including by vulgar indefinite NPs (336).

Di-phrases are subordinate to the object (325), and may be extracted to topic position (326). Their extraction as *ne*_{GEN} under DAT would replace the affectee, resulting in completely different meanings. Without *di*-phrases, *ne* takes generic readings e.g. (327) references a subject quality (indifference), rather than specific instances of feeling indifferent. In (325-326), presence of *di*-phrases indicate *ne*'s referent is communicatively relevant, requiring further specification. *Di*-phrases narrow down the broad space identified by *ne*, just as locative adjuncts intersect with existential operators (§5.4.1). As indirect references, they do *not* duplicate direct-object *ne*.

Ne is 'optional' in impersonal constructions (329-334). Russi links these to *piacere*-type (§3.3.2 for the Spanish equivalent *gustar*), however, they may be understood as impersonal transitives with inherent accusative. In (329-331), the DP is the topic, the action of caring is a comment. With *ne* (332-334), the degree of indifference is central, whilst details of the concern are secondary. Thus, unless the amount is quantified, thereby requiring *ne* (334), presence/absence of *ne* is pragmatically driven. Russi takes a similar position regarding impersonal questions (332), which neither receive interrogative intonation, nor expect informative answers. Pragmatically, 1-person pronouns convey speaker indifference towards disappointing/unpleasant situations. 2/3-person pronouns express the speaker's belief in his interlocutor's indifference toward some matter which *does* concern the speaker, often carrying derogatory overtones. Thus, (333) conveys the speaker's opinion of the addressee rather than

the topic. Without *ne*, (332) focuses on the matter (giving it importance); with *ne*, it focuses upon speaker indifference (reducing its value). Thus, whilst personal constructions focusing on subject opinion ‘require’ *ne* (SE_{DAT}+*ne*_{ACC}), impersonal constructions allow *ne*~Ø (§5.5.5 for combinations with *ci*). Many verbs follow similar patterns e.g. *importare* (335-338), for which putatively lexicalized *importarsene* is never listed, despite acting as a direct parallel (338). There is no principled means to differentiate these two sets of verbs.

Table 163

	Topic/S _H	P	D	A		
325	(Io)		me	ne _i	frego	[e _i [di lui]] I care nothing about him
326	[Di lui] _i ,		me	ne _i	frego	[e _i [e _i]]
327	Gente che		se	ne _i	frega	[e _i [Ø]] People who don’t care (about anything)
328	Chi		se	ne _i	frega	[e _i [(di lui)]]? Who cares (about him)?
329			gli	Ø _i	frega	[Ø _i [di quell’orologio]] [Something [about that watch]] matters to him
330			ti	Ø _i	frega	[Ø _i [di arrivare in orario]] [The idea [of arriving on time]] matters to you
331			mi	Ø _i	frega	[che tu arrivi in orario] _i [That you should arrive on time] matters to me
332	Che _i		mi	(ne _i)	frega	[e _i [di lui]]? What do I care about him?
333	Che		te	ne _i	frega	[e _i [Ø]]? What do you care (about it)?
334	A questi, non	glie	ne _i	frega	[niente/nulla _i [di C]]	These people care nothing about C.
335			me	ne	importa	[e _i [Ø]] It matters to me
336		non	glie	ne	importa	un cazzo ¹²⁷ He doesn’t give a shit/f*** about it!
337	Chi		se	ne	importa	? Who cares?
338	‘Me ne importa, mi sta a cuore.’ È il contrario esatto del motto fascista ‘Me ne frego’. ‘I care, I mind.’ It’s the exact opposite of the fascist motto ‘I don’t care’					

Verbs without ‘internal’ accusatives are equally compositional, but *must* express their objects (339). *Volere* takes nominal/clausal direct-, but not indirect *di*-, complements. It displays partitive usage with *ne*_{PRT} (340-345) and optional recipient datives (342-345). Like *la*_{ABS}, *ne*_{PRT} is treated as [+individuated] with direct contextual referent (340-343). If no such referent is available, [–individuated] values are sought, where its partitive nature indicates ‘part’ of a collective (‘things’). Not specifying the ‘desires’ implies something bad (a cross-linguistically

127 Cazzo: vulgar expression of disappointment/astonishment, ‘Damn!, Shit!, What the f***...!’

common euphemization strategy, Koch 2004), resulting in ‘idiomatic’ *volerne* ‘resent, desire something bad for...’(344-345). *Potere* takes clausal complements (346) pronominalized as *lo*_{PHRASAL} for [+individuated] propositions (347), or *ne*_{PRT} for the collection of ([–individuated]) propositions currently under discussion (348). When no discourse-salient referent is available, *ne*_{PRT} is interpreted as generic activity, leading to (349)’s ‘idiomatic’ reading. In *averne abbastanza* (350), *ne*_{PRT} represents *avere*’s quantified direct-object, just as *la*_{ABS} represents specific abstractions in (351-352). Again, *~sene* (342) and *~sela* (352) are purely compositional.

Table 164

	Topic/S _H	P	D	A		
339				Ø _i	voglio [due [gatti _i]]	I want two cats
340				ne _i	voglio [due [e _i]]	I want some two
341	(Dei gatti _i ,)			ne _i	voglio [[e _i]]	I want some
342			se	ne _i	voglio [[e _i]]	[+individuated] I want some for myself
343			glie	ne _i	voglio [[e _i]]	[+individuated] I want some for her
344						[–individuated] I resent her
345	Ø _i	non	me	ne _i	volere [[e _i]]	[–individuated] Don’t hold it against me
346		non		Ø _i	posso [dormire] _i con questo chiasso	I am incapable of...sleep with this noise
347	(Dormire _i ,)	non		lo _i	posso [e] _i	...it
348		non		ne _i	posso proprio più [e] _i	...it/this any more
349	Ø _i	non		ne _i	posso	I can’t go on
350				ne _i	ha avuto [abbastanza e _i [di mia moglie]]	He’d had enough of my wife
351				la _i	ha avuto [e _i vinta]	=uscire vincitore
352			se	la _i	ha avuto [e _i a male]	=rimanere offeso

Pensare’s ‘internal’ accusative (‘thoughts’) may remain unexpressed producing an intransitive quality (353), or be specified as an object (356), or proposition (354) in which case it may be pronominalized by *lo*_{PHRASAL} (355). The expressed thought may be modified (357) or expressed by *la*_{ABS} (364) if specific, or *ne*_{PRT} if indefinite (358). The thoughts may further be defined by *di*-phrases (360), extractable as *ne*_{GEN} (359), or *a*-phrases pronominalized as *ci* (361-363).

Table 165

Topic/S _H		D	A				
353			Ø _i	penso	[Ø _i] meglio con la cioccolata	I think better with hot chocolate	
354			Ø _i	penso	[che è bello] _i	I think [that it is fine]	
355			lo _i	penso	[e _i]	I think that	
356		si	Ø _i	pensa	[una bella bugia] _i	He thinks up a good lie	
357			la _i	pensa	[e _i bella]	He has a bright idea	
358			ne _i	pensa	sempre [una e _i nuova]	He's always got something new up his sleeve	
359	Cosa _i		<ne _i >	Ø _i	pensi	[e _i [<di X> _i]]	What do you think about X/it?
360				Ø _i	penso	[e _i [<di no> _i]]	I think not
361	E a M _i		ci	Ø _i	penso	[e _i tanto]	M., I think about her a lot
362	Ma tu		ci		pensi	mai al futuro?	Do you ever think about the future
363					pensa	ai fatti tuoi!	Mind your own business!
364	Ha scelto me per come la penso, non perché...					They chose me for my opinions, not because...	

Use of *ne* with these verbs calls for no special treatment; it's 'obligatory' nature (when it is required) follows from the need for transitive verbs to define their objects, whilst failure to distinguish *ne*_{ACC}~*ne*_{GEN} leads to erroneous claims of *ne* doubling *di*-phrases. Everything else follows compositionally.

5.5.5 (Ci)+Se+Ne

'Impersonal' readings available with *si*_{MID/PASS} are often difficult to distinguish from generic *si*_{IMP} e.g. *si*_{IMP} *dice che...* 'one says that...' vs. *si*_{ACC} *dice che...* 'it is said that...', both which alternate with *dice che...* 'people say...'. Combination with other clitics leads to apparent surface alternations which are treated as either lexicalized groups or evidence of clitic movement. Neither assumption is necessary.

Rendere may operate ditransitively (365) describing object (366) or subject (367, *SE*_{ACC}) transition into a state described by an accompanying adjective (*≈far diventare*). Alternatively, individuals may act as possessor/recipient (368) of the state (*Ø*_{ACC}+ADJ), with external reading 'rendered unto himself a state of X' (*si*_{DAT}) or internal reading 'becomes X_{ADJ}' (*si*_{MID}).

Table 166

	N	O	D	A	I		
365			Ø _j	Ø _i		rende un servizio _i a X _j	He renders a service to X
366				l' _i		rende e _i felice	You make her happy
367				si _i		rende e _i antipatico	He makes himself unpleasant
368			si _i	Ø _i		rende Ø _i antipatico	He becomes unpleasant
369			si			rende [Ø [conto] _{ADJ} di X]	He {becomes aware/gains understanding} of X
370			se	ne			He {realizes something/gains some understanding}...
371		ci _k	si	Ø _i		rende [Ø [conto] _{ADJ} e _k]	He realizes {it/something}...about the situation ¹²⁸
372		ci _k	se	ne			He realizes {Ø/something}...
373							Something (about it) becomes understood
374		<ci>			si	rende [Ø [conto] _{ADJ} <di X>]	One {becomes aware/gains understanding} of X
375		<ce>		ne	si		One {realizes something/gains some understanding}...
376		<ci>	Ø	Ø		rendiamo [Ø [conto] _{ADJ} <di X>]	We realize about it
377			ce	ne		rendiamo [Ø [conto] _{ADJ}]	We gain some understanding

The common phrase *rendersi conto di...* is middle (369). The subject undergoes a COS of developing (SE_{MID}) awareness (*conto*) rather than passive effect by external agent (SE_{PASS}). The object may be made explicit by *ne*_{ACC}, representing the indefinite/partial state of understanding (370). The *di*-phrase (i.e. the *content* of the growing awareness) is not a verbal argument, but subordinate to the adjective i.e. the state is one of ‘being aware of x’ as a whole. Thus *ne*≠*di*+X, as often implied in translation. It follows that it cannot be extracted as *ne*_{GEN} which would conflict with possessor *si*_{DAT}, however, it may be referenced indirectly as the current SOA through subject-oriented *ci*_{OBL}; the subject being the undergoer. (369-372) may be read with a [+SPEC] subject, or impersonally (373). With 1.PL subjects, the reflexive is *ci*_{DAT} (376-377), and *ci*_{OBL} is unavailable under RND. [-SPEC] human subjects appear as *si*_{IMP} which also takes *ci* as its dative reflexive (374, §4.6.9). Again *ci*_{OBL} is unavailable, but *ne*_{ACC} is (377, 375). The [*ci*+*se*+*ne*]~[*ci*+*ne*+*si*] alternation in (372~375) is not an example of a special placement rule, but represents distinct constructions, the meanings of which are so close that they are *treated* as equivalents.

¹²⁸ *Ce*+*se*+*ne* for some speakers. OBL clitics show -*e*/*i* dialectal variations (§6.3.2).

Table 167

	S _H	O	D	A	I		
378		m _i	se	ne		accorge	It dawns on me
379		ci	se	ne			It becomes understood between us
380	Uno		se	ne		accorge	One comes to an agreement
381			ci	ne	si		
382		Ø	si _j	Ø _i		avvale _j	[dei _{PRT} consiglio [di X]] _i He avails himself of _{PRT} advice from/about X
383		Ø	si _j	ne _i		[e _i [Ø]]	...it/some
384		<ce _k >	si _j	ne _i		avvale _j	[e _i <di ciò _k >]? ...it/some concerning that?
385			ce _j	ne _i	si _j	[e _i]	One avails oneself of _{PRT} it/some
386		ci	se _j	ne _i		frega _j	One doesn't care about it
387		ce		ne _i	si _j		
388			ce _j	ne _i		freghiamo _j	We don't care
389	%		glie	ne	si		One gives him some two (of them)
390			se	ne		regalano due	One gives some two (of them)
391	%	gli	se	ne			One gives him some two (of them) =Some two are/become given (on him)
392		ci	se	la		cava	One copes/manages
393			ce	la	si		One takes it off
394		ci	se	la		sente	One feels up to it
395			ce	la	si		One feels it

Whilst *rendere*'s state-adjective is variable, it is inherent in other verbs. *Accorger*+*se(ne)* shows similar patterns and range of meanings to *render*+*se(ne) conto* (378-381). Note that in (379), *uno* shows that *se*≠*si*_{IMP}, but must be dative. With verbs like *avvalersi*, *di*-phrases reference the source/class, whilst *si* indicates subject possessor, of partitive (382, *dei*) objects, pronominalized as *ne*_{ACC} (383) and translated 'of it' with partitive, rather than possessive, 'of'. The *di*-phrase may be indirectly referenced by *ci*_{OBL}=current SOA/topic (384). Lack of middle readings means that there is no confusion with *si*_{IMP} (385). *Fregar* (p.269) shows similar variations (386-388).

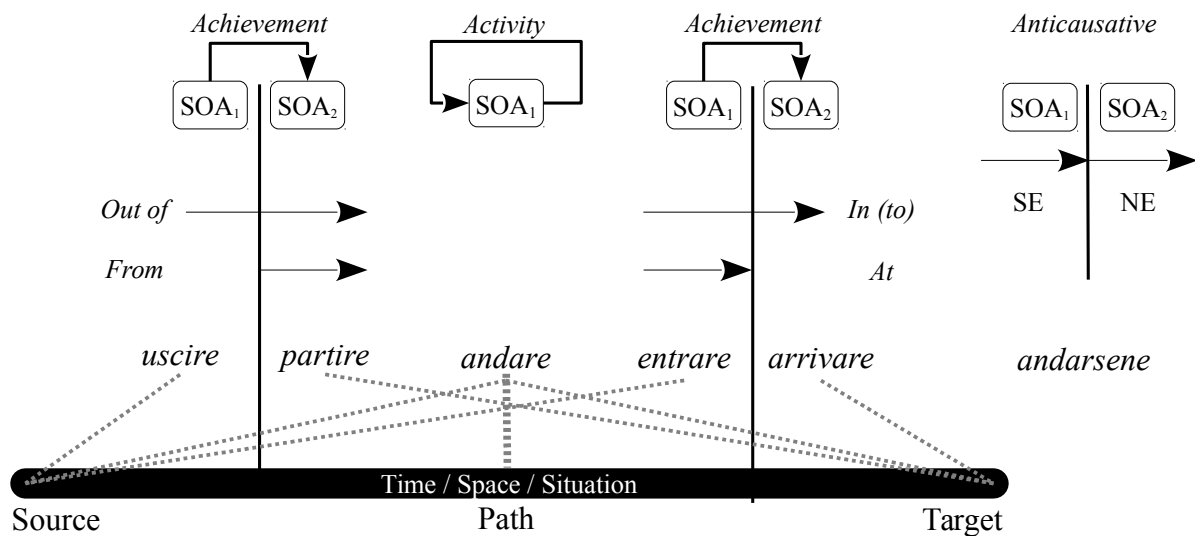
Apparent sequence variation only occurs in three clitic-clusters (Radford 1977). The alternation *cisene*~*cenesi* is not restricted to 'middle' verbs, but cases are less frequent. Generally, northern speakers accept only (389); others accept (390) and, therefore, (391).

These orders are not in free distribution nor due to optional movement, but depend on different native speakers' competences. A search of the *Libricino* corpus shows that authors (or possibly editors) are consistent e.g. Italo Calvino *ci+se+ne* vs. Franco Venturi *ce+ne+si*, reflecting dialect preferences for *expletive-it* vs. *si_{IMP}* 'impersonals' constructions. The *Dizionario Linguistico Moderne* proposes *ce+se+ne* but recommends avoiding such clusters (Gabrielli 1956:§401).

Confusion between impersonal readings of *SE_{MID}* and generic *SE_{IMP}* equally applies to [–SPEC] *la* (392-395, Lepschy & Lepschy 1984:214). Lexicalization approaches cannot cope with this degree of variation. It is only by having all items freely available that such variety could be meaningfully composed.

5.5.6 Subject-Oriented *Ne_{ABL}*

Achievement verbs of motion, inherently focus upon destination (e.g. *arrivare* highlights *SOA₂* since this is what the event has achieved) or source (e.g. *partire* highlights *SOA₁*, the achievement being one of concluding *SOA₁* and entering into a new *SOA₂*). Adjuncts may be applied to locate these events in time/space, but what the verbs describe is the achievement *SOA₁→SOA₂*. *Activity* verbs of motion e.g. *andare* do not inherently reference states of being anywhere, but the process of motion itself. They may also be associated with source-, destination- or path-oriented adjuncts. There is, however, no COS; such verbs start and end in *SOA₁*. The new *SOA₂* for achievements, and the continuing *SOA₁* for activities become the new discourse-*here*.



Availability of, and functions performed by, *se/ne* are determined by lexeme semantics. In order to create a *realization* from an *activity*, it must be delimited e.g. *run a race*, where completion is not inherent in the construction nor necessary, but may be construed from context.¹²⁹ For *activity* verb *andare* (*went*) to become an *achievement* (*up-and-went~left*), missing components for that construal must be added. SE_{ANT} presents the action as a pivotal point of change-of-state (non-motion→in-motion) located *within* the subject i.e. its focus is *up and went*. The required delimiter can logically only be source-oriented (i.e. stationary SOA₁ which is left in order to achieve the SOA₂ of movement) and is referenced by *ne*_{ABL} rather than locative phrases, since the *locus* of the achievement is *within* the person as (s)he changes state, *not* the *place* where that event occurred. Any accompanying adjuncts do *not*, therefore, double *ne*_{ABL}, but rather clarify the spatio-temporal location of the change-of-state (V+*sene*) event.

¹²⁹ Adding measures does not make predicates telic (contra some analyses); telicity can only be inferred from context e.g. 'He ran a race against her, ...but never finished' vs. '...and won first place'.

Whilst activities have unspecified duration, the interval between *ne* (starting point) and arrival at SOA₂ acts as a measure, creating a Realization as the basis for the change-of-state achievement. As an indicator of change-of-state, SE_{ANT} is inappropriate with activities, and redundant with achievements which specify, and statives which deny, it internally. Without SE, *ne* cannot reference prior states and can only be read as place or Class reference. Thus whilst (396) is read as a partitive, (397) may only be read ‘many went away’ and not ‘many of them...’ (Cardinaletti & Giusti 2006:83). There are no ‘missing’ combinations to be accounted for; semantics are reflected in syntax, and are purely compositional.

[[[V]_{ACTIVITY} +*ne*]_{REALIZATION} +SE]_{ACHIEVEMENT} : hence, *SE_{ANT}+*andare/partire/arrivare/stare*

Sene can be applied to any activity motion verb e.g. *tornarsene* (398) highlights the state from which (*ne_{ABL}*) the subject turns (i.e. changes, SE_{ANT}). Such change-of-states are often translated by ablative particles: *andare* ‘go’ vs. *andarsene* ‘go away’, *volare* ‘fly’ vs. *volarsene* ‘fly off’ (399). There is, therefore, no legitimate reason to select any specific one for lexicalization.

Source-oriented achievement verbs (e.g. *partire* ‘depart’, *uscire* ‘leave’) need not include *ne* since source=discourse-*here* is inherent e.g. *a casa* (402), but may do so in order to resume/highlight previous locations (400/403). The presence of *ne* overwrites the inherent property, indicating source individuation, to be resumed from context. As expected, such anaphoric references *are* mutually exclusive with any equally specific/individuated source adjunct in the same clause.

Table 168

396	Ne sono andati...	Some of them went...
397	Se _i ne _j sono andati via [molti Ø _i]	
398	Se ne ritornò tutto lieto a casa	He went back home all happy
399	Se ne volò	He flew off
400	Siamo arrivati a Roma; la mattina e ne _i siamo partiti la sera	We arrived in Rome in the morning and left (from there) in the evening
401	Se ne partì mentre quella insisteva nel dire che...	He took off while she protested that...
402	Stasera sto a casa, non mi va di uscire	I'm staying home tonight; I don't feel like going out
403	È entrata nello spogliatoio _i , ne _i è uscita e si è tuffata	She entered the cubicle, came out of it and dived in
404	Se n'è uscito ...senza dire nulla	She went out without saying a word.
405	...con un'altra battuta	He came out with another quip
406	Me ne sono (re)stato in silenzio	I remained silent
407	Stasera me ne sto a casa	I'll be staying {in/at home} tonight

In addition to their achievement sense, many verbs may also operate as activities in which case they may also take *-sene*. e.g. *partire*_{ACTIVITY} ‘*separate*’ focuses on ‘breaking away’ from SOA₁ rather than subsequent motion and translated ‘took (himself) off’ (401); and *uscire*_{ACTIVITY} ‘*getting out*’ (404), which may be metaphorically extended to include notions ‘escaping’ from the subject’s continuing internal state (405). Similarly, whilst usually stative, *(re)stare* may also express the activity of resisting the pull to leave SOA₁. *(Re)staresene* emphasises the achievement of staying in SOA₁ over an extended period of time, starting from *ne* (406).¹³⁰ Again, locative adjuncts merely situate this extended event (407).

130 Extension from spatial to ‘prolonged fixation in time’ follows the metaphorical cline of Heine *et al.* (1991). cf. *(re)starci* where *ci* pronominalizes static SOA/conceptual domains (§5.5.7).

Table¹³¹ 169

408	Lei se ne stava in camera	She would (continued to) stay in her room
409	Il cane se ne viveva per suo conto	The dog would (continued to) live on its own
410	Se ne andava in giro con il sorriso stampato sulle labbra	She went around with a smile fixed on her face
411	Il passerotto se ne volava nel cielo spensierato	The sparrow flew about the sky happily
412	Era contento di nuotarsene...in piscina	He was content to swim about in the pool
413	La farfalla...se ne volò via	The butterfly...flew away
414	E così quel breve pomeriggio se ne volò via	And so that brief afternoon passed quickly
415	Se ne rotolava bel bello di qua e di là	It was rolling here and there
416	vidi il teppistello corrersene...lungo la strada	I saw the little thug run off along the street
417	E tutto ormai se ne cade a pezzi	And everything is falling into pieces by now
418	Un ragazzo se ne passeggia nel giardino	A boy wanders in the garden
419	Se ne saltava da una parte all'altra	He was jumping from side to side
420	Paolo se ne dorme sul divano	P. dropped off to sleep on the sofa
421	Ecco a voi...Logan che se ne entra al ristorante	And there you have...L. entering the restaurant
422	In quel momento se ne arriva la baby sitter	At that point the babysitter arrives...
423	L.B., che se ne nacque povero in un posto infame	L.B., who had been born poor in a miserable place

As well as stative verbs (408-409), the extended time period of *ne* can be treated as a sequence of smaller activities. *Ne* points to the beginning of the sequence, defining a starting ‘measure’ (410-412). Interpretation is based on context and may be specified by further adjuncts; compare (411) with (413-414, *via*). The pattern is highly productive with all verbs of motion (415-419), but is also available with other types of verb which equally indicate an entry into an ongoing state (420). Moreover, whilst *-sene* is not found with pure achievements e.g. *raggiungere* ‘reach’ and *arrampicarsi* ‘climb up’, it is possible with activities usually associated with, but not requiring, destinations e.g. (421-423), where it is the COS which is emphasized, not final achievement.

Similar patterns are found across Romance. Sardinian’s *-sene* constructions highlight its separate forms *inke*_{ABL} vs. *inde*_{DAT/ACC} (424). Jones (1993:230-38) identifies SE’s function as focusing on the event rather than result, whilst *inke* looks back upon a prior state modified by

131 Examples taken from the ItTenTen10 corpus.

the event. The construction is often used to compensate for lack of *passato remoto* (425/426). Sardinian also has transitive constructions without SE, where *inke/nche* is also temporal/aspectual rather than locative, translated as ‘go and’ (427). Jones notes that in this use, *inke/nche* may co-occur with locative question words (428), showing that its function is not locative.

Table 170

	[SA]	[IT]
424	Lukia s-ink'est andata	Lucia se n'è andata
425	S-ink'est mortu	E' morto (exclusively verbal participle)
426	Est mortu	E' morto (verbal or adjectival)
427	Maria nk'at istrempatu sa janna	Maria (è andata e) ha sbattuto la porta
428	Ube sa balla nche fit tziu Martine?	Where on earth (lit. the bullet) was (+inke) Uncle Martin?

Contra lexicalization analyses, activity verbs alternate between \emptyset ~*sene*, because *activity*→*achievement* requires change-of-state (SE_{ANT}) and delimitation (*ne*_{ABL}) simultaneously; neither being inherent in *activity* verbs. These constructions are formed and interpreted by composition, as illustrated by their productive application to new circumstances. By way of contrast, Auger (1994:212-217) discusses several varieties of French where *s'en+voler*, *s'en+venir* etc. are becoming *se+envoler*, as shown by imperatives *Envole-toi!*, ‘Take off!’ These *are* examples of reanalysis leading to changes in the lexicon. They are both different to, and coexist with, *-sene* in French.

5.5.7 Subject-Oriented *Ci*

Subject-oriented *ci* denotes *union*¹³² with referential participants (429-430) or places (452) or propositions (433). Destination-oriented achievement verbs e.g. *arrivare* inherently denote change-of-state (motion→non-motion), hence **arrivare*+*se*_{ANT}. Such verbs do not reference

¹³² The development from *being with* (comitative) to *being in* (durative locative) a situation (conceptualized as abstract space) follows the metaphorical continuum (Heine *et al.* 1991).

prior states of motion, but focus new states of being present in/at a destination, which becomes discourse-*here*. *Ci* is not allative (motion towards), but achievement of union with a place. In this sense, *ci* provides the ‘measure’ equivalent to *ne*_{ABL} of source-oriented verbs. *Ci* may be employed to reference individuated places.¹³³ The presence of *ci* forces referential closure leading to ‘idiomatic’ interpretation when no discourse-salient place is available. Additional complements do not double *ci*, but situate the event in spatio-temporal or conceptual domains.

Stare (‘stay’, literally ‘stand’)¹³⁴ requires complements e.g. spatial-location (431), or manner adverb (432). The (c)overt complement of *starci* (‘agree with’, ‘acquiesce to’) is a proposition, with which *ci* indicates mental coincidence (433-434). *Stare con qualcuno* ‘be with somebody (romantically)’ often denotes ‘having a sexual encounter with...’, leading to colloquial idioms of *starci* ‘be easy, especially of a woman’ and *provarci* ‘attempting a sexual encounter with...’, where *ci* denotes the locus of being/participating in a situation, and euphemistic omission of the proposition invokes particular ranges of interpretations. The locus may be defined (433), but defaults to the current discourse situation/proposition (434) i.e. *ci* acts as the indirect counterpart of direct *lo*_{PHRASAL}.

Table 171

429	[Con Carlo] _i ci _i esco spesso	I often go out with Carlo
430	[Assieme a Maria] _i ci _i va sempre al cinema	She always goes to the cinema together with Maria
431	Stasera sto a casa, non mi va di uscire	I’m staying home tonight; I don’t feel like going out
432	Sto bene	I am well
433	Non ci sta. Non ci sta [a vivere una vita disperata]	He won’t go along with {it/living a desperate life}
434	Ci sto	I’m in it also

133 [–individuated] usages cannot use *vi* (Benincà 1988:177-178) which introduces distal oppositions referencing discourse-*there* (§5.2.1).

134 In central and southern varieties, *starci* acts as an existential (=esserci, §5.4.2).

Berruto (1985a), Berretta (1989), Sala-Gallini (1996) *i.a.* view *ci* in *sentire~sentirci* ‘hear~able to hear’ and *vedere~vederci* ‘see~able to see’ as a semantically empty emphatic marker. These variants, however, are neither structurally nor semantically equivalent (Russi 2008:167-8); *ci* produces contrasts (435-438). Direct-objects (437-438) are mutually exclusive with *ci*. Without *ci*, (436) must be read as if direct-objects are missing but recoverable (\emptyset_{ACC}). Although (436) can be used to refer to states of deafness/blindness, it is generally restricted to diminished ability (439), whilst *ci* is preferred for absolute inability. *Ci* is, therefore, not pleonastic; it carries stative semantic value. *Ci* denies the possibility of (c)overt direct-objects, signalling an intransitive construction focusing the SOA (*ci*). Thus, (435) is not ‘I don’t hear some/any-thing’, but ‘I exist in an ongoing state of non-hearing’, or simply ‘I can’t hear’ (incapacity). In contrast, *ci* in (440) is a simple locative anaphorically referencing the previously identified place in which the transitive event occurred.

Table 172

435	Non ci sento/vedo	I cannot hear/see ¹³⁵
436	?Non sento/vedo	?I don’t hear/see (something)
437	Non (*ci) sento nessun rumore	I don’t hear any noise
438	Non (*ci) vedo niente, è troppo buio	I don’t see anything, it’s too dark
439	Chi è presbite, infatti, vede male da vicino	Presbyopes, in fact, see badly from close up
440	Nella camera, non ci sento nessun rumore	In the room, I hear no noise

Russi considers *entrarci* (‘be involved in something’) as lexicalized. *Entrare* signifies successful completion of the subject’s physical motion into new spaces. The destination defaults to discourse-*here*, but may be anaphorically referenced through *ci* (442). *Entrarci* denotes membership of (expressed as ability to enter into) conceptual domains, also denoted by *ci*. Failure to find salient referential locations, returns ‘there’=abstract domain, often translated ‘it’. The event of entering domains, may be spatio-temporally situated by adjuncts

¹³⁵ Similarly Catalan *L’home no hi sent*, ‘The man can’t hear’.

(443), or not (444); as with physical motion (441). Since adjuncts may remain unexpressed (to be recovered from context) in both cases, accompanying PPs are not verbal arguments, but event adverbials; there is no clitic-doubling and hence no evidence for the presence of a lexicalized *ci* which has lost its “pronominal function” in addition to a real locative. Abstract place may be used with any destination-oriented motion verb e.g. *arrivare* (445-446) and *riuscire* ‘turn out, arrive at a state through one’s labour’ (447), where the mental location with which the subject becomes coincident may be made [+SPEC] by use of personal OBL_[+E] clitics (448-450). These represent a single class of verbs/constructions. There is no justification to distinguish *entrarci*. As discussed in §1.3.2, presence of a separate entry in a dictionary (e.g. *entrarci* in De Mauro 1999-2000) is not evidence for a linguistic notion of lexicalization; such entries are selected on the basis of what is considered by the lexicographer as ‘noteworthy’ or ‘helpful’ and ‘appropriate’; hence the variation in which cases appear in which dictionaries.

Table 173

441	Pinocchio entrò nel teatrino delle marionette	Pinocchio entered into the puppet theatre
442	Nel teatrino, ci entrò Pinocchio	
443	Pinocchio non c’entra con/in quella storia	Pinocchio has nothing to do with this story
444	Che c’entra?	What’s it got to do with it?
445	Non ci arrivo da solo	I can’t do (=achieve, arrive at) it on my own
446	Non <ci> arriverò mai <a capirlo>	I’ll never understand (reach understanding of) it
447	Non <ci> riesco <a farlo>	I can’t (do it) i.e. arrive at the state of...
448	Ø _i mi riesce difficile	I find it difficult (=it turns out difficult on me)
449	Ø _i mi è venuta un’idea _i	I’ve had an idea, lit. An idea has come on/to me
450	Quelle scarpe non mi entrano	Those shoes do not suit me
451	Da Roma ne arrivavano in continuazione	They were coming from Rome continuously
452	Dalla miniera ci/ne sono usciti con difficoltà	They got out of the mine with difficulty

Destination-oriented motion verbs may also reference individuated sources (451). In (452), *ne*_{ABL} may alternate with *ci*, where emphasis is on the time spent in the place (*ci*) prior

to/preparing for onset of movement, rather than initiation of movement itself (*ne*). Like (*re*)*stare*, *ci* does not represent the goal or new SOA, but continuation within the old state (discourse-*here*) without the measure provided by *ne*_{ABL}. The appropriate items are simply added to compose the desired construal.

In all these cases, *ci* references the SOA with which participants are coincident, wherever that may be in spatio-temporal or conceptual domains. It does not double anything, nor is it ever obligatory; its absence simply leads to different construals.

5.5.8 Summary

Contra Russi, *ce/ne/la* never double their complements: they have not lost pronominal status. Their presence is only ‘obligatory’ in the sense that all components of a construal must be present; without them, a different construal is formed. If the resulting verb+clitic(s) cannot be interpreted, the sentence is understood as ‘missing’ arguments and, therefore, ungrammatical. The only relationships of co-dependence are cases such as *-sene* with *activity* verbs where both are necessary to form the desired *achievement* construal. This analysis not only makes sense of examples used as evidence for lexicalization, but also for less frequent usages ‘left’ for composition, including co-existence vs. mutual exclusion of adjuncts in a single coherent approach. There is no need to consider any of these cases as lexicalized. Moreover, their range and flexibility illustrates that only compositional approaches can match the full range of data. The core meaning of the verb remains constant; whilst overall meaning is the sum of its *correctly identified* components.

5.6 Conclusions

Whilst Latin had an almost one-to-one relationship between adverb and function through morphemic concatenation, phonetic and functional syncretism during Romance's development has led to fewer forms (§2.2.2, §6.2.7). Nevertheless, their more abstract functions can be identified and their number is sufficient when associated with case to fulfil the task. Real 'confusions' do not occur due to the inner-outer interpretation (§3.3.1) and argument access (§5.1.2) processes. Remaining vagueness is infrequent and insignificant.¹³⁶

Contra lexicalist approaches (§1.3.2, §5.1.1), which see I-clitics as non-compositional items, I-clitics are not expletive, because they license ranges of abstract denotations and the variable, which they introduce, requires interpretation. Such an association is only possible if clitics are recognised as carrying a range of features including [\pm individuated], related to the clause/context through case. Each *ne/la/ci* represents a particular intersection between these properties, for which matching referents are selected in a predictable manner, and from which different meaning is composed.

As noted in the discussion of putative 'pronominal' verbs (§4.7.5), similar, sometimes identical, usages can be found across Romance, all of which developed later than proto-Romance. Lexicalization is an unlikely candidate for such parallel development over a millennium. The reason must be an underlying similarity in the meaning of the elements (clitics and verbs) and the compositional process across the language family.

¹³⁶ See §7.5.4 for an example of knowledge and active exploitation of them by speakers.

Whether composition is considered as a purely semantic process as here, or as taking place at some intermediate level of constructions as proposed by Masini (§1.3.2), they are still *composed*, not *stored* as “unanalysable chunks”. As shown, they are readily analysable, indeed it is only through this analysable status that such ‘idioms’ could have become and continue to be productive.

6 SWAPPING

Up to this point, we have argued that clitics appear as they should, and where they should, within the ‘case’ defined sequence in order to show their relationship to the verb and each other. This chapter deals with the single variation in surface sequence recognized by the model. This occurs between dative and accusative referents and is defined in terms of ‘weight’ as found in complement movement. This is a property of many different ‘close pairs’ in many different situations within a language. It is, therefore, inappropriate to provide a clitic-specific analysis.

6.1 Introduction to Swapping

In case-models, clitics are spell-outs of functional heads, underlying order being structurally determined. *Cliticization*, as combining ‘words’ into larger prosodic units is a post-lexical process influenced by prosodic environment, underlying sequence, and element properties amongst which we include *weight*. Focus upon 3-3-effects produces distorted views of the processes involved, unifying 3-3-mutations and sequence change into complex analyses requiring concepts such as clitic ‘fusion’. In our model, sequence changes are derived separately through a clitic’s *weight* relative to its syntactic partner (N~O/D~A): heavy items (indicated by superscript + e.g. *lui*⁺) move forward, unless their partner is equally heavy.¹³⁷

This is termed *swapping* to avoid confusion with syntactic ‘movement’.¹³⁸

3-3-mutations are only related to swapping if their application produces heavy clitics. Spanish (1) and Mallorcan Catalan (2) have similar 3-3-mutation rules. The difference in output

¹³⁷ Heavy constituents shifting rightwards (e.g. ‘Heavy NP shift’) is a universal functionally-motivated tendency (cf. Erteschik-Shir 1979; Arnold *et al.* 2000). See, for example, Abeillé & Godard (2000) for analysis of French complements and general word order on the basis of relative weight.

¹³⁸ This should not be confused with SE_{ANT}+OBL~OBL+SE_{MID} variations (§4.7.6) which are separate constructions where SE appears in upper or lower clitic-fields in order to express different meanings.

sequence lies in the ‘spurious’ clitic produced,¹³⁹ light *se* is static, whilst heavy *hi*⁺ advances. Although no mutation occurs in Valencian Catalan (3) or French, their sequences differ because inherently heavy *lui*⁺ must advance over light *le* (4), but not over equally heavy *en*⁺_{ACC} (5). Other factors further obscure the system. French, amongst others, has separate pro- and enclitic series with different weights, resulting in different sequences following imperatives. The overlaying of these two simpler processes creates intricate patterns with numerous ‘apparent’ exceptions used to ‘justify’ complex MCs.

Table 174

		Non 3-3-environment	3-3-environment	Rule
1	Spanish	$me_D + la_A \rightarrow me_D + la_A$	$le_D + la_A \rightarrow se_D + la_A$	$III_D + III_A \rightarrow SE + III_A$
2	Mallorcan		$li_D + la_A \rightarrow la_A + hi^+_D$	$III_D + III_A \rightarrow HI^+ + III_A$
3	Valencian		$li_D + la_A \rightarrow li_D + la_A$	} No 3-3-rule
4	French		$lui^+_D + le_A \rightarrow le_A + lui^+_D$	
5			$lui^+_D + en^+_A \rightarrow lui^+_D + en^+_A$	

Once form and relative position are established, phonetic/prosodic processes take effect, e.g. Italian *i~e* alternations. Unlike analyses which require random collections of lexicalized combinations, by separating form and sequence, such alternations can be seen to arise directly from prosodic footing (§6.3).¹⁴⁰ Catalan shows a wealth of cross-dialectal variation. With swapping explained, complexity reduces to the different 3-3-rules associated with each dialect (§6.4). Although we make no attempt to explain 3-3-rules, we take the first step by clarifying what they are responsible for, and more importantly, not.

6.1.1 The D/A~A/D Parameter

Romance clitics developed from several Latin starting points: for local-person clitics $MIHI_{DAT} \sim ME(UM)_{ACC} > m\bar{i}_{DAT} \sim m\check{e}_{ACC}$ etc.; a range of ‘heavy’ adverbial sources reduced in form to produce non-personal clitics e.g. $HINC > ci/hi$, $IBI > vi/bi$, $IPSE > se$; and 3-person

¹³⁹ The term “spurious” here refers to any clitic which appears “unexpectedly” in place of another.

¹⁴⁰ For the complexities of Romanian prosody, see (§7.4.4).

clitics developed from the distal adjective/WP ILLU(M) etc., the result of which may be ‘heavy’ or ‘light’ depending on the language. Since datives/locatives tended to derive from ‘heavier’ sources, Proto-Romance showed predominantly D/A-order.

During Romance’s development, phonetic erosion in unstressed environments produced quantitative and qualitative vowel reduction such that personal dative and accusative clitics converged upon single forms e.g. Italian *mĩ*, Spanish *mě*. Romanian, having preserved dative case (but not vowel length), is the only modern language to consistently distinguish *mi*_{DAT}~*me*_{ACC}, and even here syncretism amongst plurals is customary (§7.4.4). ‘Weight’, derived from such early morphological/phonetic distinctions, and experienced as sequence variation, was stored for each form as part of the grammaticalization process of WPs into modern clitics. Middle French provides another and later example of this process in the development of its independent enclitic series (§6.10).

The inadequacy of a language-wide parameter is shown by historical developments. Although swapping generally decreases as clitics progress to ‘light’ with heavy non-personal accusatives *en/ac* and adverbials *y/en* lasting longest, clitics (often in groups) change at different rates, thereby changing the overall pattern of sequences and ‘exceptions’ e.g. Provençal 1/2+3-combinations *lo*_A+*me*_D became *me*_D+*lou*_A during XV^c, but D/A-order 3+3-combinations do not appear until XVII^c (Wanner 1974:164). Processes such as phonetic reduction and paradigm uniformity tend towards weight equalization and hence D/A- (i.e. underlying case-) order. This trajectory has completed in some languages (Spanish and Portuguese show consistent D/A-order from the earliest records, (Menéndez Pidal 1904:304), but many languages retain some heavy elements producing mixed patterns (e.g. Italian). In

some, A/D order has been actively increased e.g. the dialect of Roergat (§6.5). Thus, although the overall process is $A/D > D/A$, it cannot be expressed as a parameter. The process is not binary, but emergent, and based on the granularity of individual (or group) weights.

6.1.2 Spell-Out

Although original *weights* stored as results of grammaticalization reflected contemporary morphology, no modern language employs vowel length as a morphological property, or accent patterns based on heavy/light syllables, although some retain consonant length. Heavy items may be reflected in phonetics e.g. Italian *ci* where the palatal consonant is geminated, or multi-morphemic appearance e.g. French *lu-i*, but not necessarily. In many cases, the same form shows different syntactic behaviours, indicating its use to represent (related) ranges of underlying feature sets, but always shows the same swapping properties (e.g. *hi*, §6.4.1). Each generation of children learns clitic *weights* by positive experience of each surface-form's behaviour in multiple combinations/environments. They associate *weight* with form, not with putative (and silent) underlying morphemic structures, the nature/organization of which cannot be ascertained from experience.

Morphological/syntactic 'rules' have the effect of changing feature-sets associated with case positions. Whatever the feature-set's source (underlying or mutated), the clitic which matches that feature-set, for that case, is spelled-out. When that results in surface-forms of different relative weights, swapping may occur, followed by language-specific phonetic/footing rules. The overall result is a complex set of ordered pairs which may appear to require 'lexicalization', but are in fact entirely transparent, and more importantly, learnable.

6.1.3 Chapter Outline

§6.2 briefly introduces the origin of spurious 3-3-forms, which will appear repeatedly throughout the chapter. §6.3-6.7 focus on proclitic sequences, divided between languages generally taken to represent the D/A (Italian/Catalan) vs. A/D (Occitan/Aragonese) dichotomy, highlighting not only the inadequacy of such descriptions, but also the range of unnecessary theoretical complications which follow from such concepts. §6.8-6.9 focus on languages which show enclitic changes in form, order and/or stress. It is shown that by separating out swapping from prosody, such variations follow the same logic as proclisis. Contra analyses based on WPs, post-verbal sequence variation is determined by (potentially weight-bearing) allomorph selection which is shown to be independently necessary. §6.10 takes French as a case study. The complex range of phenomena found both pre- and post-verbally across dialects/registers are examined and found to follow naturally from the above. Finally, §6.11.1 considers ‘feature transfer’, the only remaining case of 3-3-context feature ‘arithmetic’ found in the literature, providing a speculative (given the limited data) solution which follows directly from our argumentation and provides a better fit to the empirical data.

6.2 The Nature of Spurious 3-3

This section takes Italian (which we argue, contra previous analyses, does have a 3-3-rule) as an example and then compares the arguments presented with similar developments across Romance.

6.2.1 Orthography and Structure

Benincà & Cinque (1993:2325) suggest that orthographic variation *me+lo+V* vs. *V+melo* and *glielo+V* reflect separated vs. conjoined underlying structures.¹⁴¹ The conjoined sequences are

¹⁴¹ Until recently Italian high-school grammars condemned *gli* as 3.DAT.PL clitic, recommending post-verbal *loro* (e.g. Marinucci 1996); *gli presta loro il libro*, ‘he lends them the book’, but such use of *loro* has only marginal status among speakers of Standard Italian (Cordin & Calabrese 2001:551). Conversely, use of *gli* has been widespread throughout Italian’s history (Serianni 1988:213), even in written contexts (Hall 1960).

not, however, phonological words, since word-internal processes do not apply (Vogel 2009). In enclisis, orthographic conjunction serves to separate/distinguish clitics from WPs which also follow verbs, whilst there is no such motivation preceding verbs as shown by the fact that Italo Calvino used to write *glie lo* whilst proclitic *melo* is a common childhood mistake (Cardinaletti 2008:65). Thus, writing merely reveals language-group orthographic conventions, *not* structure.

For theory-internal reasons, several proposals separate clusters containing third- vs. local-person datives, regardless of pro-/enclisis. Thus, *glielo* forms “a unique clitic constituent at the structural level” (Laenzlinger 1993:253) or an “amalgamation”, best analysed in morphology (Heggie & Ordóñez 2005:26). Both cluster types, however, display identical *surface* properties in syntax (e.g. non-separation under clitic-climbing), prosody (e.g. secondary stress placement) and phonology (e.g. initial clitic $i \rightarrow e$). While *glie*-forms stand-out as products of 3-3-rules, there is no *a priori* reason to treat them differently, merely theory-bound ones. We proceed on the basis that all clitics are equal and independent, regardless of their orthography.

6.2.2 Morphemic Structure and Markedness

Unlike most Romance languages, Italian distinguishes singleton clitic DAT.SG.M *gli* [ʎi] vs. DAT.SG.F *le*. Both, however, become *glie*- [ʎe] in 3-3-combinations. Cardinaletti (2008:64) considers 3.ACC clitics and DAT.SG.F *le* to be bi-morphemic ($l+e$). Unlike *-i* in *gli* and *-e* in other clitics, *-e* in *le* cannot delete before vowel-initial verbs: *Gli/Gl’/Le/*L’ ho aperto la porta*. Cardinaletti argues that, unlike epenthetic *-i*, class-marker *-e* is morphologically too complex to be the first element in “single-word” clusters, leading to replacement by simplex,

We, therefore, treat *gli*_{DAT.PL} as a full member of the clitic lexicon.

hence less-marked, *gli*. However, if non-deletion proved bi-morphemic status, 3.ACC clitics (*l+o/a/e/i*) should also prohibit vowel deletion, but their reduction is commonplace. The approach also ignores cross-linguistic evidence. In Spanish, *le(s)+lo→selo* <Old Spanish *gelo*, i.e. both simplex *le* and complex *le+s* are replaced by simplex *se* (identical to the reflexive). If simplex→simplex is possible, it is not bi-morphemic status (which may be independently true) that determines change. Moreover, spurious-*se* derives from de-palatalization of Old Spanish *ge* [ʒe].¹⁴² During its use, *ge* had no other function in the language and was, therefore, more marked than what it was replacing. Thus, markedness cannot be the source of [ʎi]/[le]→[ʎe].

6.2.3 3-3-Rules

The key observation is that *glie-* [ʎe] only occurs in 3-3-contexts; beyond DAT/ACC, where *gli/le* are OBL, such changes do not occur (6). Moreover, 3-3-product *glie-* is distinct from its sources. Unlike Spanish and Romanian, Italian disallows dative-doubling (7), except with 3-3-combinations (8-9, Benincà 1988:137). *Glie-* is not doubling the dative complement, but performing a different function, regardless of its *gli/le* source (Benincà & Poletto 2005:232). Due to the PCC, only *glie+ne/lo/la/le/li* arise. Under our model, this is a 3-3-effect (3.DAT+3.ACC→3.OTHER+3.ACC) whereby datives are replaced by a non-dative (impersonal locative)¹⁴³ which happens to look like *gli*_{DAT} in Italian and the reflexive in Modern (but not Old) Spanish. Reduction of gender/number contrast derives from this process, with no structural implications. Many dialects of Catalan show a further development, where *hi* (= *ci*) has spread to 3-3-contexts. Thus, for Italian *ci*_{LOC}~*ci*_{IMP}~*glie*₃₋₃~*gli*_{DAT}, Central Catalan shows *hi*_{LOC}~*hi*_{IMP}~*hi*₃₋₃~*li*_{DAT} (Bonet 1991:211-212).

¹⁴² Schmidely (1978) for detailed developments.

¹⁴³ Manzini & Savoia (2002) and Řezáč (2010) argue that 3.DAT is syntactically a kind of locative clitic.

Table 175

6	A Maria _k , di zucchero, nel caffè _i , le _k ce _i ne _i metto sempre troppo _i	I put too much _i there _i for her _k
7	(*Gli) ho regalato il libro a Mario	I have given ...the book... to him (Mario)
8	Glie=l' ho regalato a Mario	...it...
9	Glie=ne ho regalati due a Mario	...some two...

6.2.4 Motivation/Nature of OTHER

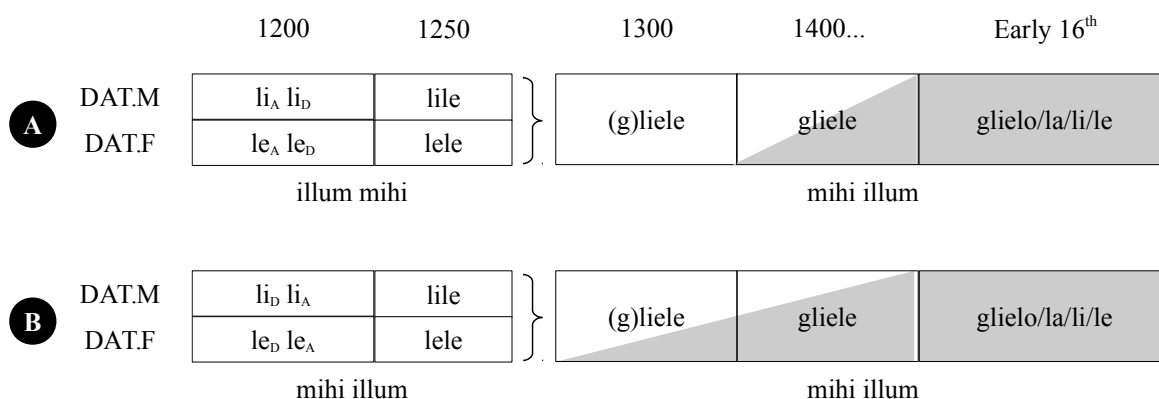
RND restricts clitics of equal (e.g. **mi+mi*) and overlapping (e.g. **mi+ci*) identity for local-persons which are speech-act unique. 3.DAT+3.ACC clitics, however, may have distinct referents. Where these are referentially unique due to reflexivity e.g. *se+le/le+si* (where [+R]=SUBJ~[-R]≠SUBJ), no change occurs. When referentially equivalent e.g. *gli+lo* (where either clitic might refer to either participant, even though their referents may be distinguishable from context and/or by the accusative clitic), mutation (at morphological/syntactic level) is required. In sentences with single clitic and complement, the latter is highlighted, backgrounding clitic referents. With two clitics, the action is highlighted and both participants are backgrounded. 3-3-rules reflect the relationship between backgrounded participants i.e. focused ACC vs. 'other'.

In Italian/Spanish, this process only appears to check dative person. It maintains gender/number information about focused ACC, whilst reducing the secondary participant to generalized 'other'. Catalan dialects show a range of 3-3-rules (§6.4), many of which produce different 'dative' outputs depending on input number and/or reduce ACC to 'generic' *ho/Ø*. French appears to have no 3-3-rules, but may convert datives to *y_{LOC}* in some circumstances, and frequently drops accusatives in clusters. There are, therefore, many possible resolutions to the situation, but in each case, it is the *referent's* underlying properties (reflexivity/number/person) which determine whether 'mutation' occurs and the final output, *not* notions of markedness or sub-structure.

3.OTHER is mutually exclusive with datives/locatives without performing dative functions whilst lacking gender/number. It is convenient to place it in what is arguably its historical source position; [III,LOC,-SPEC]. This is notably not accessible directly in any language (vi_{LOC}/ci_{LOC} must be referential), but could surface as the result of feature-changing processes.

6.2.5 Development of *Gli*

When pronouns became clitics, bisyllabic DAT.PL *loro* was problematic. Its slot in the clitic lexicon remained empty, forcing use of post-verbal *loro*. *Glie* developed in clusters, and was later abstracted to stand alone as *gli*.



Wanner (1974:162) claims that Old Italian 3-3-clusters were characterized by “special morphological manifestations [...] *lili* for a masculine dative, and as *lele* for a feminine dative”, assuming that the first syllable represents 3.ACC.M/F.SG/PL, whilst the second indicates masculine (*li*) and feminine (*le*) datives. (A) illustrates Wanner’s view of its historical development, which leaves *lili*→*lile* unexplained and contrary to the general process of raising *e*→*i* in weak positions.

Data from the OVI indicate that *glie*-clusters with ACC agreement appeared much earlier (Russi 2008). Given that dative and accusative are identical in the earliest phase, analysing the

sequence as A/D rather than D/A is based solely on presumption of language-wide A/D→D/A. The development is better explained as (B).

In Old Italian, homophonous 3.DAT and 3.ACC.M.PL *li* contextually palatalized, $li\#V \rightarrow lj\#V \rightarrow \Lambda\#V$, creating a *li~gli* [li]~[ʎi] alternation affecting both clitics. Gradually these allomorphs specialized: *gli_{DAT}~li_{ACC}*.¹⁴⁴ This process co-existed with an optional phonological rule whereby final unstressed *e*→*i*, producing alternations such as *avante~avanti* ‘before’ (Rohlf 1966:178, also §6.3.2). The alternation *gli~glie* [ʎi]~[ʎe] arises naturally, therefore, *iff* the first element of the pair [ʎe] alternating with isolated [ʎi] was the dative.

Table 176

10	se Egli me la concede	If He (God) grants it to me (data from Pescarini 2013)
11	che [...] voi la mi concediate	That you may grant it to me
12	lo 'mperadore lo si trasse di sotto	The emperor took it out from below himself
13	e assai ne _i gli piacquero e _i	Many were pleasing to him
14	ché gli ne potrebbe troppo di mal seguire	Because it could cause him too much misfortune
15	che gli le demo p(er) una inpossta	That we gave them to him as a tax

[ʎele]’s final *-e* might be expected to raise to *-i*. Its invariability indicates that it is a particular, not accidental, form. We propose that *le* derives from a 3-3-rule, reducing ACC to a common form (with an underlying, rather than epenthetic, vowel), just as Old French ACC→Ø, and Catalan dialects ACC→*ho*/Ø (see below).

For local-person pronouns, whilst the earliest records exhibit A/D-order, both orders were acceptable by the 1300s (10-11).¹⁴⁵ Notably, reflexives pattern with local-person (12). As noted above, only 3-3-contexts of referentially equal partners require conversion to OTHER. This was never so for *(g)li~(g)l(e)*. The only cases of *ne+gli* are different constructions e.g.

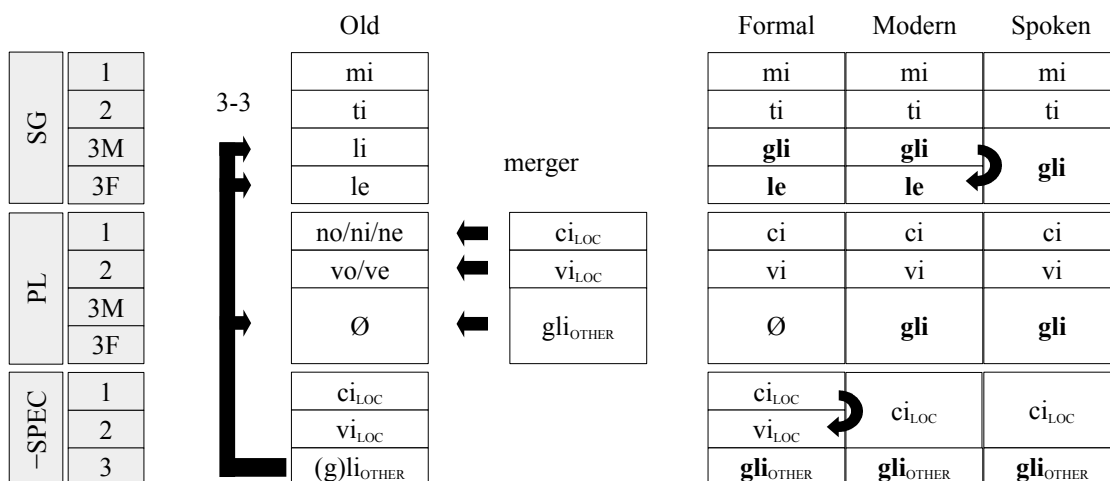
¹⁴⁴ Even when homophonous, these forms never collapsed into a single syncretic item, but remained distinct. Conversely, *glie* spread to 3.ACC.M.SG/PL and 3.DAT.SG/PL in Arce (Pescarini 2007).

¹⁴⁵ Aski & Russi (2010) for a quantitative survey and pragmatic-based account of this alternation.

where *ne*_{NOM} precedes *gli*_{OBL} in a purely intransitive construction (13). On the contrary, invariable *ne*_{ACC} (14) and *le*_{ACC} (15) always follow *gli*_{DAT}, producing the contrast: *le*=definite~*ne*=indefinite.¹⁴⁶ *Gli*'s appearance indicates a further stage of reanalysis (see below) leading to increased use of specified ACC in *glielo/la/le/li/ne*. Thus, whilst early A/D~D/A is found with local-person clitics, 3-3-combinations were D/A from at least 1250. At no time, need these clusters be considered lexicalized units.

6.2.6 Generalisation of *Gli*

Singleton~cluster variation *li*+Ø~*gli*+*lo* (cf. *mi*+Ø~*me*+*lo*) leads to *gli* being abstracted as 3.DAT.M.SG outside of clusters, contrasting with already present 3.DAT.F.SG *le*. Since the 3-rule replaces both singular and (non-existent) plurals, *gli* is also abstracted to DAT.PL, explaining why no DAT.F.PL variant developed. At this point, *gli* represents DAT.SG.M and DAT.PL.M/F, which when in 3-3-contexts, is replaced by the same product of *li* as before, i.e. it looks as if nothing has changed, except in the case of feminine singular. Nevertheless, its doubling behaviour (§6.2.3) shows that it has.



146 Piobbico (Marche, Manzini & Savoia 2005) shows a similar pattern with 3.DAT *i* in isolation and 3-3-contexts, whilst 3.ACC mutates: [+DEF]→*li*, [-DEF]→*ni*.

In another process, $no(s)_{1PL}/vo(s)_{2PL}/\emptyset_{3PL}$ (<Latin NOS/VOS, only rarely attested in XIII^c) were replaced by $ci_{LOC}/vi_{LOC}/gli_{LOC}$ (Rohlf 1968; Cardinaletti & Egerland 2010). If the placement of OTHER is correct, gli_{OTHER} moves into DAT.PL, confirming the non-purist trend. Indeed, reanalysis continues. For most speakers, vi_{LOC} has become redundant, whilst 3.DAT.F “*le* survives only in very careful or formal speech when speakers want to maintain what is perceived to be a higher standard.” (Russi 2008:92) notes that many, who believe that they use *loro* and *le* correctly, in reality use *gli* quite consistently in unguarded use. Thus, in practice, *gli* fills 3.DAT.M/F.SG/PL and 3.OTHER.

In contrast, Old Spanish had already developed gender-less DAT.PL *le+s*. The OTHER replacement for 3.DAT.M/F.SG/PL was *ge* ([ʒe]<[ʎe]<li) coexisting and alternating with *se* in reflexive contexts. With the loss of palatal fricatives, *ge* [ʒe]>*se* [se], producing the notorious spurious-*se* rule. The only reason that the Italian spurious-*glie* rule goes un-remarked is that it looks like DAT.SG.M (but not DAT.SG.F), rather than a clearly spurious clitic.

6.2.7 3-3-Rules Across Romance

3-3-rules are often discussed in terms of avoidance of two identical sounds (16), however in most cases this cannot be their motivation (17). §2.2.2 showed purely phonetic developments causing syncretism e.g. *ni* and *inde*. This section shows how such changes combine with structural developments.

Table 177			
16		17	
Spanish	*le+lo→se+lo	Sarroch	*ddi+ndi→si+ndi
Italian	*le+lo→glie+lo	Italian	*le+ne→gliene
Napoli	*le+lo→nce+lo	Napoli	*le+ne→nce+ne
Grotaglie	*li+lo→ni+lo	Barceloní	*li+en→n'i

Along with dative ILLI(S), Latin also possessed locative ILLI (>Italian *lì* ‘there’). It is necessary to account for three locatives: proximal INCE, medial IBI and distal ILLI. Excepting Italian, these converged at the surface and semantic level as generalised locatives, most often derived from INCE (spatio-temporal proximity being extended to discourse-*here*), although Sardinian generalised IBI>bi. We propose that distal ILLI provided the basis for 3-3-forms. Table 178 shows that as its surface-forms phonetically develop, they often converged with partitives, reflexives, and 3.DAT. In languages where INCE/IBI spread to other locative positions including ILLI, 3-3-forms may converge with locatives. In a final stage, 3.DAT~3-3-form alternations may lead to the 3-3-form replacing 3.DAT outside of 3-3-contexts, and the loss of number/gender distinctions e.g. Italian *glie* (§6.2.5).

Table 178

ILLI	>	[li]	>	[ʎe]	>	[ne]	>	[ne]
				[ʎi]	>	[ʒe]	>	[se]
					>	[ni]	>	[ni]
					>	[ʒi]	>	[si]
INCE	>	[ɲdʒi]	>	[dʒi]	>			
	>	[ɲʃi]	>	[ʃi] ci				
	>	[ɲʃe]	>	[ʃe] ce				
	>	[ɲge]	>	[ge] g(h)e				
IBI	>	bi	>	(h)i/y				
	>	vi						

Rohlf's (1968) notes processes of nasalization where *gli* becomes *gni* (Firenze, Lucca, Capoliveri), *gne* (Sinalunga, Cortana), *ni* (Pisa, Santa Maria de Guidice), or *ne* (on Elba). As Rohlf's notes *ni/ne* must develop from *gni*, not *inde*. In Lecce (Pescarini 2007), *ni* replaces *gli* in clusters and isolation, contrasting with *nde* and *nci*. Several southern dialects (Manzini & Savoia 2005), have replaced 3.DAT forms with *ne*<*gne*. Nociglia maintains 3.DAT in isolation (18), but *nɛ* in 3-3-contexts (14),¹⁴⁷ whilst Nocara has *nə* in both contexts (21-22),

147 Similarly Rocca Imperiale (*i+i→ni* (<*ni+i*), Manzini & Savoia 2005), where it is syncretic with partitive and 1.PL pronouns, Castrovillari (*li+lu→ni+lu*, (Loporcaro 1995), Spinazzola and Grottaglie (Melillo 1981).

resulting in the possibility of two surface-identical, but positionally and featurally distinct, forms co-existing (23). In Celle di Bulgheria (24-27), 3-3-contexts produce *ɲɔʒi*, which is a pre-nasalised palatal (<INCE) also used as the locative (27), whilst INDE>*ni*.

Table 179

18	li _i Ø _i 'dajɛ 'kwistu _i	He gives this _i to him _j	Nociglia (Apulia)
19	lu/la/li/lɛ 'viʃu	I see him/her/them	
20	nɛ _i lu/la/li/lɛ _i 'dajɛ	He gives it/them _i to him _j	
21	nə _i Ø _i 'ða stu 'kundə _i	He gives this _i to him _j	Nocara (Calabria)
22	n _i u/a/i _i 'ðaðə ¹⁴⁸	He gives it/them _i to him _j	
23	nə _i nə _i 'ða d' du _i :jə _i	He gives (some) two _i to him _j	
24	li _i Ø _i 'danu 'kistu _i	They give this _i to him _j	Celle di Bulgheria
25	lu/la _i 'viðinu	They see him/her _i	(Campania)
26	ɲɔʒi _j lu _i 'danu	They give it _i to him _j	
27	ɲɔʒi _j Ø _i 'mittu 'kistu _i	I put this _i there _j	
28	a _{SCL} g _i Ø _i 'dag kwas-'kɛ _i	I give this _i to him _j	Modena (Emilia)
29	a _{SCL} g _i al/la/i/li _i 'dag	I give it/them _i to him _j	
30	a _{SCL} g _i la _i 'mɛt	I put it _i there _j	
31	nende+bi+lu	Telling=him=it	Ossi (Sardinian)
32	ɖɖi/ɖɖizi _i Ø _i a k'kustu _i	He gives this _i to him/them _j	Làconi (Sardinian)
33	ndi _i ɖɖi _i a d' duaza _i	He gives some two _i to him _j	
34	si _i ɖɖu _i 'aða	He gives it _i to him/them _j	

	ILLI	ILLIS	ILLORUM	INCE	IBI	OTHER (ILLI)	PRT
Old Spanish	le	les		(y)		[ʒe] ge ₃₋₃	Ø
Modern Spanish				Ø		[se] se ₃₋₃	
Old Italian	le/li	Ø	loro	ci	vi	[ʎe] glie ₃₋₃	ne
Modern Italian	le/gli	gli		ci			
Nociglia	li					nɛ ₃₋₃	nɛ
Nocara	nə					nə ₃₋₃	nə
Celle di Bulgheria	li			ɲɔʒi		ɲɔʒi ₃₋₃	
Lecce	ni	ni		nɛ		ni ₃₋₃	nde
Arce	glie			ce		ce ₃₋₃	
Napoli	le			ce		nɛ ₃₋₃	
Poggio Imperiale	i			cə		cə ₃₋₃	nə
Cantanzaro	ⁿ ce			ⁿ ce		ⁿ ce ₃₋₃	
Modena	g			g		g	
Ossi (Sardinian)	li	lis		bi		bi ₃₋₃	
Làconi (Sardinian)	ɖɖi	ɖɖizi		bi		si ₃₋₃	

In other dialects, INCE spread to 3.OTHER e.g. Poggio Imperiale (S. Italy, Manzini & Savoia

148 The *ða*~*ðaðə* alternation is phonologically determined.

2005:135-138) $i+u \rightarrow cə+u$. Arce has *glie* in isolation, but *ce* in combinations and as locative. Napoli uses an earlier form for clusters, $le+^l o/ne \rightarrow nce_{3-3}+lo/ne$, whilst its locative has continued development to *ce*. Like Old Spanish ge_{3-3} , nce_{3-3} has no other function in the language. In Catanzaro, *ce* is used in all situations. INCE often voiced e.g. Modena (28-30, Manzini & Savoia 2002), where locative *g* fulfils all functions. In Sardinian, IBI>bi spreads to all locative positions and generally converts DAT $\rightarrow bi_{3-3}$ (31). Làconi (Sardinian, Manzini & Savoia 2005) has 3.DAT $qd\dot{i}^+/qd\dot{iz}i^+$ in isolation (32), and in combination with ACC_[-DEF] (33), but *si* with ACC_[+DEF] (34). Rohlfs (1966:156) notes *si* as 3.DAT in parts of Calabria e.g. Ardore: *si parlau* (gli parlò) even outside of 3-3-contexts. Also Manzini & Savoia (2004:46) for S. Agata del Bianco (Calabria), which they consider equivalent to spurious-*se*.

The sections below illustrate a wide variety of triggers for 3-3-replacements and effects of weight.

6.3 Italian

This section aims to show that, once weight has been recognised, Italian's combinatorial sequences are as transparent as those of Spanish. Unlike most Romance languages, Italian shows alternations in singleton~cluster vowel realization, sometimes used to infer lexicalized pairs. §6.3.2 offers an alternative explanation based solely upon structures already posited.

6.3.1 Basic Patterns

In ditransitives with animate recipients, the PCC restricts combinations to inanimate 3.ACC which are all heavy, resulting in no swapping regardless of DAT's weight (35-37). For spatial destinations, ACC is unrestricted and heavy ci^+_{LOC} (but not light vi^-_{LOC}) advances over light 1/2.ACC (39), but not heavy 3.ACC (38,42).

Pairs of personal *ci/vi* and locative *ci/vi* are incompatible (43). They may combine where one is locative and the other personal, but due to *ci*'s weight, produce the same surface-sequence (44-45). In combination with other personal pronouns, *ci*⁺, but not *vi*⁻, advances (39-40). Since *ne*⁺_{ACC} is heavy, no movement occurs (57). Unlike some French dialects (§5.2.2), pairs of *ne*'s are ungrammatical (Cinque 1995:195). OBL participants¹⁴⁹ may be added (41-42), creating similar surface-sequences to (39-41) but with different meanings. Pescarini (2007) notes some speakers' use of *ci*_{OBL} (46-48), where emphasis is laid upon receipt by *mi* at a place, rather than arrival at a place for *mi*'s benefit (41-42). There are, therefore, two potential meanings for *mi+ci* (and, for some speakers, two for *ci+mi*).¹⁵⁰ Clearly, no person-sequencing model can explain such variations.

Table 180

D/A	Ø	mi	ti	ci	vi	si ⁺ _{ACC}	(lo/la/li/le) ⁺	ne ⁺
Ø	Ø+Ø	Ø+mi	Ø+ti	Ø+ci	Ø+vi	Ø+si	Ø+lo/la/li/le	Ø+ne
me	mi+Ø						me+lo ⁺	me+ne ⁺
te	ti+Ø						te+lo ⁺	te+ne ⁺
ce	ci+Ø						ce+lo ⁺	ce+ne ⁺
ve	vi+Ø						ve+lo ⁺	ve+ne ⁺
se	si+Ø						se+lo ⁺	se+ne ⁺
gli	gli+Ø						glie+lo ⁺	glie+ne ⁺
le	le+Ø							
ne ⁺ _{DAT}	ne+Ø						%ne ⁺ +lo ⁺ ¹⁵¹	*ne ⁺ +ne ⁺
ve _{LOC}	vi+Ø	vi+mi	vi+ti	vi+ci		vi+si ⁺	ve+lo ⁺	ve+ne ⁺
ce ⁺ _{LOC}	ci+Ø	mi+ci ⁺	ti+ci ⁺		vi+ci ⁺	ci ⁺ +si ⁺	ce ⁺ +lo ⁺	ce ⁺ +ne ⁺

Italian *si*_{IMP} follows all clitics, as indicated by its position in the model (§2.1). Since each *si* has a different syntactic position, surface forms with other clitics may appear to alternate. In

149 These are given in standard form. Some speakers use *-e* forms for OBL (§6.3.2).

150 Maiden & Robustelli (2000:§6.4) notes that native acceptability judgements are “by no means clear-cut where ‘locative’ and first and second person clitics are concerned”, advising language learners is to avoid such combinations.

151 e.g. <ne> *lo tolse* <da li>. Said to be archaic or dialect restricted, but often found on the internet.

(52-53), past-participle agreement shows that *le* is accusative in both cases, accompanied by *si*_{IMP} (52) and *si*_{DAT} (53), as reflected in translation. With *avvalersi*, *si*_{DAT} represents the subject taking possession of indefinite/partitive (54, *dei*) objects, pronominalized as *ne*_{ACC} (55) and translated ‘of it’ with partitive, not possessive, ‘of’. The *di*-phrase (i.e. the object’s source/class) may be indirectly referenced by *ci*_{OBL}=current SOA/topic (56), which may then be confused with the DAT_[+R] of *si*_{IMP} (57). The *ce+se+ne~ce+ne+si* alternation (56~57) is not swapping¹⁵² but two distinct constructions. Clearly, no template-based model can explain such variations. As discussed in (§3.4.5), cases such as (58) are not PCC breaches, but OBL+SE_{DAT/MID}.

Table 181

		O	[D	A	X	[I	V _i]]		
35	G _i		mi _k	Ø _j			porta _i	il libro _j	G _i brings {the book _j /it _j /two _j } to me _k
36			me _k	lo _j				e _j	
37			me _k	ne _j				due e _j	
38			ce _k	lo _j			porta _i		G _i brings it _j there _k
39			mi _j	ci ⁺ _k					G _i brings me _j there _k
40			vi	mi					
41		mi	ci _k	Ø _j				il libro _j	G _i brings {the book _j /it _j } there _k for me
42		mi	ce _k	lo _j					
43			<*ci>	ci			accompagna		He takes us there
44			vi	ci ⁺					...you there
45			vi	ci					...us there
46		%ci		mi			porta _i		G _i brings me _j there _k
47		%ci	mi	Ø _j				il libro _j	G _i brings {the book/it} to me there
48		%ci	me	lo _j				e _j	
49	Ø _i		glie _k	lo _j		si _i	porta _i	e _j	One _i brings it _j to him _k
50			glie _k	ne		si _i		[due e _j]	...some two _j to him _k
51			mi _j	ci ⁺ _k		si _i			...me _j there _k
52			le			si	è vendute	bene	One sold them well
53			se	le					He sold them to himself
54			si	Ø _j			avvale	[dei _{PRT} consiglio [di X]] _j	He avails himself of _{PRT} advice about X
55			se	ne _j					...it
56		ci	se	ne _j				e _j	...that about X
57			ce	ne _j		si _i			One avails oneself of it
58	Ø _i	mi	si _i				avvicinò	un mendicante _i	A beggar came up to me

152 Compare with (51), where swapping takes place behind *si*_{IMP}.

6.3.2 Prosody

Vowel change in clusters has been explained as a historical process resulting in lexicalized clusters (Gerlach 2002), or a synchronic phonological lowering rule (e.g. Cinque 1995:194). Since this rule is inapplicable in identical phonological contexts e.g. *mi/*me lava* ‘he washes me’, it must be expressed as cluster-internal (59, Kaisse 1985). Kayne (2000:154) notes that some speakers ‘allow vowel change’ in triplets on OBL clitics separated from the sonorant (*A Mario, lo zucchero, nel caffè, non glie ce l’ho messo*, ‘I did not put it there for him’), but not in similar D/A-clusters (**Me ci⁺ metterà* ‘he will put me there’). Rather than ‘vowel change’, we argue that OBL clitics in such dialects simply have underlying *-e* whilst the ‘rule’ only applies to syntactic pairs (60). Cardinaletti (2008:62) notes dialects with both clitics in *-e*: *Me ce metterà*, i.e. where most clitics end underlyingly in *-e*, and no rules apply.

Table 182

59	$[(\text{CL.DAT} \dots \text{i})][\text{CL.ACC} \dots \text{i})] \rightarrow (\text{CL.DAT} \dots \text{e})[\text{CL.ACC} \dots \text{i})]/$	}	____[coronal sonorant]	(ce lo) porta
60	$[(\text{CL.NOM} \dots \text{i})][\text{CL.OBL} \dots \text{i})] \rightarrow (\text{CL.NOM} \dots \text{e})[\text{CL.OBL} \dots \text{i})]/$			(se ne) va

There is, however, no phonetic basis for the lowering process. A more insightful answer, requiring neither lexicalized clusters nor spurious phonological rules, is to see the change as the residue of a prosodic rule once pervasive in Italian which has been re-analysed as part of clitic-specific prosody.

In early Florentine, verbal pronouns were WPs (separate bi-moraic feet) with many clusters taking A/D-order. When D/A-order appeared (indicating reanalysis as clitics), dative vowels in clitic-pairs changed to *-e* with few exceptions (Melander 1929). During the same period (Rohlf 1966:178), an optional phonological rule whereby final unstressed *e* \rightarrow *i*, gave rise to alternations such as *avante~avanti* ‘before’, and eventually resulted in separation of clitics

and stressed pronouns which retain etymological *-e* (*mi*_{CLITIC} vs. *me*_{PRONOUN}). Clitics in isolation or cluster final (weak position) were subject to raising (*e*→*i*), whilst initial clitics (strong position) escaped the rule. Thus, *i*~*e* alternations became diagnostic of prosodic status: *-e* in foot-heads, *-i* in weak/extra-metrical positions. For clitic vs. stressed pronouns, this distinction was lexicalized, remaining long after the phonological rule became moribund. Within clitic-clusters, it was re-analysed as part of the clitic-field's prosody: etymological *mi*_{DAT}~*me*_{ACC}>prosodic *mi*_{WEAK}~*me*_{STRONG}. For Standard Italian, OBL are underlying *-i*, *lo/la/le/li/ne* retain their etymological vowel, whilst vowels of other clitics alternate based on position. Paradigm uniformity may lead to simplification e.g. OBL clitics (Kayne's dialect) or all clitics except *lo/la/le/li* (Cardinaletti's dialect) end in *-e*.

Each syntactic pair may form a foot, inducing *e..i* sequences. Items separated by syntax e.g. *mi*+*si*_{IMP}, *mi*_{OBL}+(*sela*) do not form feet at this level. Since PW phonology (e.g. s-voicing) is not found, we assume that such feet are independent elements within CG.¹⁵³ Re-syllabification at higher levels of prosody runs sets together (including verbs and negatives), but *e..i* patterns remain fixed within the feet, which phrasal re-syllabification must respect. Evidence for such feet, and the distinction between the two classes, comes from pronunciation where strong positions are phonetically lengthened e.g. [me:lo] but *[mi:si], and the ability to truncate (i.e. squash into a single bi-moraic foot e.g. *ce+lo*→*cel*, *ce+ne*→*cen*, but **mis* <*mi+si*) in poetry e.g. Old Italian *s'ella è dessa, più non **mel** celate* (Pescarini 2007).

In clitic triplets, OBL remains extra-metrical and surfaces with *-i*. Heavy dative clitics advance over light accusatives. The resulting pair does not form a foot, and both vowels surface as *-i*. Under phrasal re-syllabification, the palatal of *ci*⁺ which is always treated as long

¹⁵³ The relationship between prosodic words (PW) and clitic groups (CG) is developed in §6.8.4-§6.8.5.

inter-vocally, prevents *mi*'s vowel lengthening; $ci^+ + mi \rightarrow [mic.ci]$. Similar patterns occur in the upper clitic-field e.g. *si~se+ne*, *ci~ce+ne*. Unfortunately, use of personal OBL with SE_{ANT} e.g. Spanish $se_{ANT} me_{OBL} murió$ does not exist in Italian (§3.3), so it is impossible to test the effect of *me/te/gli* on *se/i*, however, $SE_{NOM} + ACC$ constructions would seem to imply that two extra-metrical items may also form a foot at the higher level of phrasal re-syllabification/footing e.g. $(se_{NOM} lo_{ACC}) mangió tutto$, 'he ate it all up'.

6.3.3 Locatives

Vi_{LOC} raises difficulties since Modern Italian barely uses it (§5.2.1) and acceptability judgements are even weaker than for ci_{LOC} . Our searches provided only examples of *vi* in isolation or requiring vi_{OBL} . Whereas ci_{LOC} represents discourse-*here*, vi_{LOC} displaces time/place, representing events from a different viewpoint and so is limited to situations of opposition (61-62). Its only common usage is $vi_{LOC} + ci_{1.PL}$ replacing $*ci_{LOC} + ci_{ACC}$, where it could equally be OBL, which would explain its *-i*. In $ci^+_{LOC} + vi_{2.PL} \rightarrow vi_{2.PL} + ci^+_{LOC}$, they do not form a pair, and so remain unchanged.

Most cases of $ci + si$, are subject-oriented ci_{OBL} (63-64) or $ci_{REF.DAT} + si_{IMP}$ (65). Combination with SE_{PASS} is unacceptable regardless of animacy (66-67) since passives do not accept DAT even when locative. Even as SE_{ACC} , usage appears to be questionable (68). It is possible to read this as (69), where *ci* is once again OBL. Some speakers, however, do accept the paradigm (70-75). Swapping indicates that these are D/A pairs. Weight correctly predicts sequence, but not the vowel i.e. $[(ce:.si)]$ might be expected. It might be that $ci + si \rightarrow [cis.si]$ (cf. $[mic.ci]$) helping to explain why pairs ending in *si* cannot reduce in poetry, however, there do not appear to be any phonetic studies to support or deny this.

Table 183

		O	(D	A)	X	I		
61	\emptyset_i	vi_i		mi			vede	He sees you there
62		si					oppose	He opposed (himself) there
63	\emptyset_i	ci_i	se	<le _i >			lava <le mani _i >	He washes {them/his hands} (there)
64	In quel ristorante _i ,					si	mangia bene	In that restaurant, one eats well there
65			ci	\emptyset_i		si	mette il libro _i	One puts {the book/it} there
66			ci	si _{PASS}				*The book is put there
67			ci	si _{PASS}				*He is put there
68			ci	si _{REF}			mette lui	??He puts himself there
69		ci	si					?He puts his self there
70	(io)		mi	ci ⁺			abituerò	[mic.ci]
71	(tu)		ti	ci ⁺			abituerai	[tic.ci]
72	(lui)		ci ⁺	si ⁺			abituerà	*[ce:.si]
73	(noi)		vi	ci ⁺			abitueremo	[vic.ci]
74	(voi)		vi	ci ⁺			abituerete	[vic.ci]
75	(loro)		ci ⁺	si ⁺			abitureranno	*[ce:.si]

The analysis confirms that *glie+lo/la/li/le/ne* are no different to *me/te/se/ce/ve+lo/la/li/le/ne* (§6.2.1-6.2.2). The only combination that could require lexicalization is *ci+si* in this very particular usage, and low-frequency collocations are not good candidates for such a process.

6.3.4 Syntactic Approaches?

Pescarini (2013) presents much the same data as evidence for a syntactic approach. Building on the Linear Correspondence Axiom (Kayne 1994:19-21), the WP→clitic evolution resulted in changes in syntactic configuration, from split sequences (i.e. clitics occupying different, although adjacent, A..D positions), to true clusters (i.e. single complex heads where dative clitics left-adjoin to accusatives). This distinction is manifested, not only in A/D~D/A-order changes, but also in absence of 3-3-mutations in languages which retain A/D-order (split configuration) e.g. French.

Table 184

		A		D			A		D	
76	Old French	Je	[le	[te	comande	Et cils	[le	[li	dient	
77	Modern French	Je	[te+le	[comande	Et ils	[le	[lui	dient	
78	Quebec French	Je	[te+le	[comande	Et ils	[lui+le	[dient	
'I command it to you'					'and they tell it to him'					

Whilst such syntactic explanations are attractive (also Ordóñez 2002), they provide no means of identifying: which items changed in which language; why WPs such as Italian *loro* do not operate similarly; or its timing, given that Quebec French *lui* became light after the period of change-over (§6.10.3). Moreover, many dialects with A/D-order in 3-3-contexts do mutate datives (e.g. Corsican, §6.8.2), whilst the theory has nothing to say about similar effects in the upper clitic-field. The complication introduced by this particular implementation, including ACC dominating DAT, seem unwarranted. Indeed, on theory-internal grounds, (Kayne 2008) now assumes that “sequences of clitics never form a constituent”. We, therefore, retain our simpler approach. It provides greater coverage and facilitates cross-linguistic comparison, including French (§6.10).

6.4 Catalan

Catalan¹⁵⁴ displays a vast range of dialect variations. This section considers eight of the most studied in order to illustrate that the differences can be expressed by minor changes in their clitic lexicon without resorting to complex mechanisms or processes.

3.ACC gender vowels (M. *o/u*, F. *a/ə*) may be prosodically suppressed. *Ho/hi* (not found in all dialects) never delete, but may form diphthongs with preceding vowels or phrase-initially. *Vos/nos/mos* may lose final-*s* (Eivissa: *mu ne dunaràn*), or even reduce to *s* (València:

¹⁵⁴ Examples from Perea (2012, itself a digest of Alcover (1916), Alcover & Moll (1929-1933).

a/nem's-en). Full/syllabic forms are found in prosodically strong post-verbal positions, however, some speakers maintain pre-verbal full forms before consonants and in fossilized expressions e.g. *Déu vos guard!*, *Quant ne vols?* Epenthetic ə (emboldened) is common: *vuz-**e** n'aneu*, especially with (e)ls: *aquestes taules elz-**e** les vendré* (S. Llorenç de Cerdans).

Table 185

AL	Alta Llitera (Ribagorçà dialects)			Sistac i Vicén 1993; Bonet 2002	
NVS	Non-Valencian Standard			de Borja Moll 1968:171–2; Bonet 1993, 2002	
MO	Monòver			Colomina i Castanyer 1985; Todolí 1992; Segura i Llopes 1998	
MJ	Marina Baixa			Colomina i Castanyer 1985, 1991; Todolí 1992	
MA	Mallorcan varieties			de Borja Moll 1968, 1980; Bonet 1993, 2002	
VS	Valencian Standard			Todolí 1992; Bonet 1993, 2002	
BAC	Baix Camp varieties			Bonet 2002	
BC	Barceloní			Bonet 1995, 2002	

			Non-syllabic		Syllabic		‘•’ ‘ə’, ‘i’ suppressed vowel potential epenthesis		
			Proclitic	Enclitic	Proclitic	Enclitic			
I	S		m		em	me	[mə]/[m•]/[^ə m•]		
	P			ns	ens	nos	[nos]/[n•s]/[^ə n•s]/[^ə n•s ^ə]/[mos]/[m•s]/[^ə m•s]/[^ə m•s ^ə]/[mo]/[s]		
II	S		t		et	te	[tə]/[t•]/[^ə t•]		
	P			(us)	us	vos	[vos ^ə]/[us ^ə]/[vo]/[s]		
III	A	M	S		l		el	lo	[lo]/[l•]/[^ə l•]
			P			ls	els	los	[los]/[l•s]/[^ə l•s]/[^ə l•s ^ə]/[əs]
		F	S		l		la		[la]/[l•]
			P				les		[ləs]/[l•s]/[^ə l•s]/[^ə l•s ^ə]/[əs]
	D	S				li		[li]	
		P			ls	els	los	[lis]/[l•s]/[^ə l•s]/[^ə l•s ^ə]/[^ə l•s ⁱ]	
	R		s		es	se	[^ə s ^ə]		
	A		n		en	ne	[^ə n ^ə]		
					hi		[i]/[əj]		
	N				ho		[o]/[w]		

Although DAT.PL is often syncretic with ACC.M.PL, some dialects have *lis*, whilst others have the much disputed [(ə)lzi] (§6.4.2). [MA] frequently uses *le(s)* for masculine datives and accusatives. [BC] systematically removes gender-markings in all combinations, i.e. all plural cases and genders surface as *‘ls* ±epenthetic *e/i*. Many western varieties have 3.PL /ez/ in addition to /l(e)z/ (Todolí 1992:143). Typically, /ez/ is pre-verbal and /l(e)z/ post-verbal, however, Tàrbena Catalan allows both in pre-verbal position; other varieties use /ez/ for ACC

and DAT simultaneously, [ez-es] (Bonet 2002:956). /l(e)z/~/ez/ alternations do not interact with opacity e.g. [MJ] 3.PL.DAT+3.ACC surfaces as [liz-o]/[əz-o]. Since the same alternation is found with definite articles and undergo similar modifications (Colomina i Castanyer 1985: 161-63), and /ez/ may appear in isolation, the alternation cannot derive from clitic interaction; /ez/ is a selectable allomorph. Sequence is effected by neither form (e.g. post-verbal *los* for '*ls*, or pre-verbal *ez* for *lez*) nor stress e.g. when Balearic dialects displace stress to final clitics (§6.8.10).

6.4.1 Sequence-Variation

Before XVI^c, Catalan followed mostly A/D-order, now preserved solely in Mallorca (Alcover 1916). In [VS]/[MO]/[MJ] all combinations are D/A with localised 3-3-rules.¹⁵⁵ The remaining dialects also have heavy *hi*⁺_{LOC}/*en*⁺_{ABL} where *hi*⁺ may also be the result of 3-3-rules.¹⁵⁶ Created or underlying, heavy clitics advance except against other heavy *en*⁺_{ACC}/*ho*⁺. [MA] has heavy local-person datives, causing these to also advance. 3-3-rules are sensitive to dative number.

As a standard variety, [VS] is considered artificial (Todolí 1992; Bonet 2002): all combinations surface transparently. [NVS]'s clitic system derives from older stages of Catalan (Casanova Herrero 1990). DAT.SG→*hi*⁺, whilst DAT.PL surface transparently, like [VS]. In [AL], which has 3.PL.DAT *lis* in isolation and /a/ as feminine marker, the 3-3-rule affects both singular and plural whilst the advanced *hi*⁺ forms diphthongs with the open vowels of accusatives, a tendency found in all dialects, but so consistent in [AL] as to be formalized in its description. In all three, ACC clitics are identifiable by their gender markings.

¹⁵⁵ Rare 2+1 combinations are not DAT+ACC, but OBL+DAT/ACC (Chapter 7).

¹⁵⁶ Western dialects ([MJ]/[VS]/[MO]) retain subject-oriented *hi* e.g. *no hi veu/sent* ≈Italian *senitirci/vederci* (§5.5.7), but not the object-oriented *hi* under discussion.

In [MO], DAT→*se* in 3-3-contexts, probably under Spanish influence (Casanova Herrero 1990; Todolí 1992). Clitic order is D/A as shown by gender-marked accusatives. 3.PL allomorphs are not syncretic, but pre-verbal /es/ and post-verbal /los/, hence [se-s] (Segura i Llopes 1998:61-63). When pre-verbal DAT.3PLs combine with *ho/en*, they may surface as [ez] rather than *s'*. It is unclear from the description and limited data whether this represents optional epenthesis preceding heavy clitics or a more complex 3-3-rule.

In addition to *hi*⁺_{LOC}/*en*⁺_{ABL}, [MA] has heavy 1/2.DAT clitics which advance unless the accusative is equally heavy *ho*⁺_{ACC}/*en*⁺_{ACC} e.g. *dóna-me*⁺*-les*→*dóna-les-me*, but *torna-mos*⁺*-ho*⁺→*torna-mos-ho*. In all three, DAT.SG→*hi*⁺. In [MA], with DAT.PL, ACC→*ho*, sometimes shows dative gender [əlz-o]~[lez-o].¹⁵⁷ [BAC]/[BC] have distinct DAT.PL /lzi/ (not reflected in the orthography), but unlike [BAC], [BC] suppresses gender-marking vowels.¹⁵⁸ In [BAC]/[BC], DAT.PL triggers ‘generic’ accusative Ø (vs. [MA/MJ] *ho*). This results in identical surface-forms from multiple sources e.g. [lzi]</ls+hi/ or /lsi+Ø/. If [BAC]/[BC] had selected *ho* as generic accusative, the difference would be clear.

In [MJ] (which also has *lis*), DAT.PL causes ACC→*ho* [o]/[w]. With DAT.SG, matters are disputed. Before XVIII^c, [MJ] followed [BC]’s pattern (*li+lo/la*→*lo·y/la·y*) including accusative specificity constraints whereby *ho*→’*l* (§6.4.2), producing *lo·y*. The new pattern emerged following development of transparent *li·u*<*li+ho* (Colomina i Castanyer 1991:62). Todolí (1992) sees this as *ho* spreading to all combinations, innovating plural-marked [wz] by analogy with /lz/, but this doesn’t explain its limitation to DAT.3.SG. Nor can [w(z)] be an exponent of ACC number since this would require it to also appear with DAT.PL.

¹⁵⁷ [MA] *les* may be used as DAT.M/F, and even ACC.M (de Borja Moll 1968:170).

¹⁵⁸ [BC] drops feminine-markers in 3-3-combinations, but not with other persons (*Les sabates, me les donarà la Teresa*) and masculine-markers in all combinations (*comprar-lo(s)* vs. *compra(r)-me’l(s)!*).

Table 186

79	en ⁺ _{ACC}	ho ⁺	el _{SG.M}	l(a) _{SG.F}	els _{PL.M}	les _{PL.F}	hi ⁺ _{LOC}	en ⁺ _{ABL}	
General	et	te-n	t ⁺ ho	te-l	te-la	te ⁺ ls	te-les	t-hi ⁺	t-en ⁺
	us	vos/us-(e)n	vos/us-ho	vos/us-(e)l vos/us-lo	vos/us-la	vos/us-els vos/us-los	vos/us-les	vos/us-hi ⁺	vos/us-en ⁺
	em	me-n	m-ho	me-l	me-la	me-ls	me-les	m-hi ⁺	m-en ⁺
	ens	(e)n(o)s-(e)n	(e)ns-ho nos-ho	(e)n(o)s- (e)l (e)n(o)s-lo	(e)n(o)s-la	(e)n(o)s-els (e)n(o)s-los	(e)n(o)s-les	(e)n(o)s-hi ⁺	(e)n(o)s-en ⁺
	es	se-n	s-ho	se-l	se-la	se-ls	se-les	s-hi ⁺	s-en ⁺
VS	els	(e)l(o)s-(e)n	(e)l(o)s-ho	(e)l(o)s-(e)l	(e)l(o)s-la	(e)l(o)s-els	(e)l(o)s-les	/	/
	li	li-n	li-ho	li-l	li-la	li-ls	li-les	/	/
MO	els	es-en/se-n	es-u/s-o	se-l(o)	se-la	se-(lo)s	se-(le)s	D→SE	D→SE
	li	s-en	s-u/s-o	se-l(o)	se-la	se-(lo)s	se-(le)s	D→SE	D→SE
MJ	els	e(l)z/liz-en	e(l)z/liz-ho	e(l)z/liz-ho	e(l)z/liz-ho	e(l)z/liz-ho	e(l)z/liz-ho	A→HO	/
	li	li-n	li-w	li-w	li-w	li-wz	li-wz	/	/
hi _{LOC} en _{ABL}		Ø+en	hi+Ø	l-hi ⁺	la-hi ⁺	(e)l(o)s-hi ⁺	(e)l(e)s-hi ⁺	HI ⁺ =>	EN ⁺ =>
		Ø+en	en+Ø	l-en ⁺	l(a)-en ⁺	(e)l(o)s-en ⁺	(e)l(e)s-en ⁺	HI ⁺ =>	EN ⁺ =>
NVS	els	els-en	⁺ ls+Ø	⁺ ls-l	⁺ ls-lə	⁺ ls-ls	⁺ ls-ləs	/	/
	li	l ⁺ en	li+Ø	l ⁺ -hi ⁺	lə-hi ⁺	⁺ ls-hi ⁺	ləs-hi ⁺	D→HI ⁺	D→HI ⁺
AL	els	els-en	⁺ ls+Ø	lo-j/je	la-j/j(e)	l(o)s-i/je	las-i/je	D→HI ⁺	D→HI ⁺
	li	l ⁺ en	li+Ø	l-i/lo-j	la-j	l(o)s-i/je	las-i/je	D→HI ⁺	D→HI ⁺
BAC	els	els-en	(ə)l(u)zi+Ø	(ə)l(u)zi+Ø	(ə)l(u)zi+Ø	(ə)l(u)zi+Ø	ləzi+Ø	A→Ø	D→HI ⁺
	li	l ⁺ en	li+Ø	l ⁺ hi ⁺	l ⁺ hi ⁺	(ə)l(u)z-hi ⁺	ləz-hi ⁺	A→Ø	D→HI ⁺
BC	els	(e)lsi+Ø	(e)lsi+Ø	(e)lsi+Ø	(e)lsi+Ø	(e)lsi+Ø	(e)lsi+Ø	A→Ø	D→HI ⁺
	li	n ⁺ hi	li+Ø	l ⁺ hi ⁺	l ⁺ hi ⁺	⁺ ls-hi ⁺	ls-hi ⁺	A→Ø	D→HI ⁺
MA	els	els-en	els+Ø	⁺ ls-ho	⁺ l(e)s-ho	⁺ ls-ho	⁺ l(e)s-ho	A→HO	D→HI ⁺
	li	l ⁺ en	li+Ø	l-hi ⁺	l-hi ⁺	⁺ ls-hi ⁺	⁺ ls-hi ⁺	A→HO	D→HI ⁺
	te ⁺	te-n(e)	t-ho	(e)l-te ⁺ lo ⁺ t ⁺	(e)lz(e)-te ⁺ les/los-te ⁺	la-te ⁺	(e)lz(e)-te ⁺ les-te ⁺	3-3-RULES	
	vos ⁺	vos-en vos-ne	vos-ho	(e)l-vos ⁺ lo-vos ⁺	(e)lz(e)-vos ⁺ les/los-vos ⁺	la-vos ⁺	(e)lze-vos ⁺ les-vos ⁺	Swapping	
	me ⁺	me-n(e)	m-ho	(e)l-me ⁺ lo-m ⁺	(e)lz(e)-me ⁺ los-me ⁺	la-me ⁺	(e)lz(e)-me ⁺ les-me ⁺		
	mos ⁺	mos-en mos-ne	mos-ho	(e)l-mos ⁺ lo-mos ⁺	(e)lz(e)-mos ⁺ les/los-mos ⁺	la-mos ⁺	(e)lz(e)-mos ⁺ les-mos ⁺		

Bonet (2002:957) discounts *l*-vocalization as [w(z)]'s source, however, this is the understanding of grammars (Fabra 1956) and language-wide dialect studies (Alcover 1916, Alcover & Moll 1929-1933). It explains when it appears, /li+l/→[liw], /li+ls/→[liwz], and how it emerged; loss of *hi* in these circumstances triggered emergence of transparent *li+ho*

and *li'l(z)* which became vocalized [*liw(z)*]. Alcover provides several cases of ‘*l(s)→[w(z)]*’ from Marina Baixa itself e.g. *els llibres no puc comprar-li-us*; but no cases with feminine nouns. Neighbouring areas provide definite cases of *l*-vocalization /*la/→[ua]* e.g. *torna-li-ua!* (Simat de la Valldigna). Most examples display ‘standard’ forms. Given the paucity of data and regional tendency to sporadic *l*-vocalization, we follow ‘traditional’ analyses.

Thus, 3-3-rules may include spurious datives, *se⁻/hi⁺* and/or ‘generic’ accusatives, *ho⁺/Ø* which, since *ho* is heavy, has no effect on sequence. Contra de Borja Moll (1980:29–30), 3-3-combinations do not present “una varietat de solucions gairebé anàrquica.” Whilst phonetic processes such as *l*-vocalization obscure matters, the overall pattern is readily discernible. Nevertheless, the emboldened items warrant elaboration.

6.4.2 Complex Forms

DAT.3.PL has two forms (Bonet 1991, 1995; Viaplana 1980): normative *els* [əlz] of high registers and some North-Western dialects; and *els hi* [əlzi], the colloquial form of Central/North-Eastern Catalonia, apparently combining *els+hi*. Martín (2012) believes that DAT.3.SG [*li*] should also be understood (as sometimes written) as *l’hi*. Along with [əlzəni] and [ni], [*i*]’s ‘random’ appearance has generated numerous morphological analyses.¹⁵⁹

Bonet (1993), Harris (1996), Solà-Pujols (1998) *i.a.* treat [*i*] as a dative case morpheme within the structure [*l_{DEFINITE}+Ø/z_{PLURAL}+i_{DATIVE}*] but, since [*i*] does not appear in local-person datives, its morphemic status seems questionable. For Martín (2012), [*i*] is a deictic morpheme, where datives are complexes subsuming accusatives; [[*l+Ø/z*]_{ACC}+*i*]_{DAT}. However, availability of post-verbal *los_{DAT}* (**losi*) in these dialects and [*l_{DEFINITE}+i/e/o_{VOWEL}+z_{PLURAL}*] in others, show that, despite historical origins, modern forms are *lexical* items which have drifted so far that

¹⁵⁹ Examples from Bonet (1991, 1993, 1995a, 1995b).

no sub-structure can be reliably demonstrated. Fortunately, simpler explanations are available.

Bonet (1993:91-92) presents the data such that two singulars produce [li] (80), but if either is plural (81-83), [əlzi] appears. From this, Bonet argues for clitic ‘fusion’, similar to accounts of American Spanish dialects, which putatively show DAT-ACC feature interchange (Harris & Halle 2005). This explanation cannot hold for Catalan since [əlzi] also appears in isolation (85), where no accusative clitic can source such operations (84/85~80/81). In fact, Mascaró & Rigau (2002:10) state explicitly that [əlzi] is only available when accusative clitics are absent. Despite recognising that ACC-ellipsis is common across Romance, including “restricted areas of the Catalan speaking domain”, Bonet rejects it because [əlzi]’s plural-marker “has to come from the accusative clitic”, but this merely leaves (85)’s [əlzi] unexplained.

Table 187

80	El llibre, al nen, [li] dono demà	I will give	...the book to the boy... tomorrow
81	El llibre, als nens, [əlzi]...		...the book to the boys...
82	Els llibres, al nen, [əlzi]...		...the books to the boy...
83	Els llibres, als nens, [əlzi]...		...the books to the boys...
84	[li] dono el llibre		...him/her the book...
85	[əlzi] dono el llibre		...them the book...

As illustrated in Table 188, in [BC] 3-3-contexts, DAT.SG→*hi*⁺_{PCC} and advances, whilst DAT.PL sees ACC→∅. In [MA], DAT.PL triggers ACC→*ho*. In [MJ], DAT.PL also causes ACC→*ho*, but DAT.SG does not trigger conversion to *hi*⁺_{PCC}. [əlzi] appears as an open-syllable allomorph of *els*_{DAT} (regardless of ACC-ellipsis or absence due to the presence of a complement), performing the same disambiguatory function as *lis*_{DAT} in [MJ]. Outside of 3-3-combinations, heavy *hi*⁺_{LOC} also advances, producing *l’hi/els’hi* as a separate process. [VS] has neither 3-3-rules nor *hi* so that surface forms are transparent, and [əlzi] is not produced (except as free variants by some speakers). Contra Bonet, the plural-marker of [əlzi] is DAT’s

plurality which triggered ACC-ellipsis. In all dialects, once the 3-3-rule's bipartite nature is recognised, [əlzi]/[əlzo], *els 'hi* [əlzi], *l'hi* [li], and *li* [li] appear as expected.¹⁶⁰

Table 188

D A		[BC]			[MA]			[MJ]			[VS]		
S	S	li→hi ⁺ _{PCC}	l+hi ⁺ _{PCC}	[li]	li→hi ⁺ _{PCC}	l+hi ⁺ _{PCC}	[li]	li	li+l	[liw] ¹⁶¹	li	[li'l]	3-3- Product
	P	li→hi ⁺ _{PCC}	els+hi ⁺ _{PCC}	[əlzi]	li→hi ⁺ _{PCC}	els+hi ⁺ _{PCC}	[əlzi]	li	li+ls	[liwz]	li	[ləls]	
P	S	A→Ø	els+Ø	[əlzi]	A→ho	els+ho	[əlzo]	A→ho	els+ho	[əlzo]	els	[lsəl]	Generic ACC
	P	A→Ø	els+Ø	[əlzi]	A→ho	els+ho	[əlzo]	A→ho	els+ho	[əlzo]	els	[əls əls]	
S	Ø		li+Ø	[li]		li+Ø	[li]		li+Ø	[li]	li+Ø	[li]	Open- syllable
	P		els+Ø	[əlzi]		els+Ø	[əlz]		lis+Ø	[liz]	els+Ø	[əlz]	
Ø	S		Ø+'l	[l]		Ø+'l	[l]		Ø+'l	[l]	Ø+'l	[l]	
	P		Ø+els	[əlz]		Ø+els	[əlz]		Ø+els	[əlz]	Ø+els	[əlz]	
L	S		l+hi ⁺ _{LOC}	[li]		l+hi ⁺ _{LOC}	[li]		l+hi ⁺ _{LOC}	[li]			Locative
	P		els+hi ⁺ _{LOC}	[əlzi]		els+hi ⁺ _{LOC}	[əlzi]		els+hi ⁺ _{LOC}	[əlzi]			

As illustrated in (86), [li]/[ni] have several sources. Since [VS] lacks *hi*⁺_{LOC}/*en*⁺_{ABL} (Bonet 1991:73) and 3-3-rules, all combinations surface transparently.¹⁶² In most dialects (represented by [NVS]), ditransitive objects must be specific, and hence represented by 'l; ACC_[−SPEC] surfaces as Ø. Thus DAT+*ho* never appears; rather underlying DAT+*l*/Ø surfaces as appropriate to each dialect. In many cases, adverbial clitics are unexpressed giving the same result as [VS]. Similarly *hi/en+en* do not surface; specific *en*_{ACC} (≈'l) is required and DAT is dropped; non-specific *en*_{ACC}→Ø. Again, these underlying forms surface as appropriate to each dialect (79). For some speakers, *en*⁺_{ABL} triggers 3-3-rules producing *l'hi*. In [BC], this is always so. Furthermore, [BC]'s *en*_{ACC} is light resulting in *hi*⁺ (LOC or 3-3) advancing over it, producing *n'hi* [ni].¹⁶³ The unexplained forms are [əlzəni]/[əlzin], which Bonet states are acceptable variants for some speakers of these dialects.

160 Pescarini (2007:295)'s generalization of 'datives mutate but accusatives drop' requires revision. ACC-ellipsis must be seen as *substitution* by Ø_{ACC}, matching *ho*_{ACC}. In both cases, the substitute is [3.ACC,−SPEC]; variation derives from whether that slot in each dialect's lexicon holds Ø or *ho*.

161 As noted earlier, we take these to be cases of *l*-vocalization.

162 *Li+ho* is only found in "*el Reine de València viu*" (Alcover 1916).

163 Unlike French, such changes derive from inherent weight alone, not pre-/post-verbal position: *en*_{ABL} is heavy (*si tu l'hi poses, ell l'en traurà; treu-l'en tu*), and [BC]'s *en*_{ACC} is light (*n'hi posaré una; posa-n'hi una*).

Table 189

86	[VS]		[NVS]						[BC]	
	ho _[±SPEC]		ho _[+SPEC] →‘l		ho _[-SPEC] →∅		en ⁺ _{ACC}		en ⁺ _{ACC}	
li _{DAT}	li+ho	[liw]	li+lo	l’hi ⁺ _{PCC} [li]	li+∅	[li]	li+en	[l’en]	hi ⁺ +en	n’hi ⁺ _{PCC} [ni]
els _{DAT}	els+ho	[əlzo]	els+lo	els-l [ls]	els+∅	[əlz]	els+en	[əlzən]	els ⁺ +en	els+∅ [əlzi]
								[əlzin]		[əlzəni]
hi ⁺ _{LOC}	∅+ho	[ho]	hi ⁺ +lo	l’hi ⁺ _{LOC} [li]	hi ⁺ +∅	[hi]	hi+en	∅+[ən]	hi ⁺ +en	n’hi ⁺ _{LOC} [ni]
en ⁺ _{ABL}	∅+en	[ən]	en ⁺ +lo	l’en ⁺ _{ABL} [lən]	en ⁺ +∅	[en]	en+en	∅+[ən]	en ⁺ +en	n’hi ⁺ _{PCC} [ni]
				l’hi ⁺ _{PCC} [li]						
∅	∅+ho	[ho]	∅+lo	’l [l]	∅	∅	∅+en	[ən]	∅+en	[ən]
DAT.SG→hi ⁺										

6.4.3 [(ə)lz(ə)ni]/[(ə)lzin]

Taking OBL into account increases available combinations with [i]. (87-89) show *l(s)'hi* alongside OBL. Whilst 3.OBL alone produce *li* and *els* (90-91), combination with pronominalized locatives produce further cases of *l'hi* [li] (92) and *els hi* [(ə)lzi] (93). Whilst *els*_{OBL}+*en*_{ACC} produces [(ə)lzən] (94), pronominalization of *hi* creates [(ə)lzan] in [BC] where *en*_{ACC} is light allowing *hi*⁺ to advance (95), or [(ə)lzin] in dialects where *en*_{ACC} is itself heavy (96). For many speakers, *hi* is simply dropped leaving [(ə)lzən] (94).

Fabra (1956) warns against *els n'hi* for *els en*. (98) is acceptable, because *els* is OBL; its interpretation forced by presence of three clitics. If, however, *els* is DAT i.e. recipient/possessor, it clashes with equally DAT *hi*. The presence of [i] in [(ə)lz(ə)ni] indicates the advancement of underlying *hi* forcing DAT to be erroneously read as OBL (98). Thus (98) cannot be used to mean (97). Fabra's warning, however, implies that speakers are want to do so. Indeed spoken language often makes use of 'pleonastic' *hi*. (99) can also be expressed as (100) where 'there' is recognised as a topical participant/situation i.e. *hi* is an impersonal dative used to distance speaker and recipient. Both clitics may combine (101), where *els* is OBL, a third party affected by the telling event, but not necessarily the recipient. As with

many uses of OBL, grammarians disapprove, and such forms are avoided in formal registers.

Contra Bonet, $[(ə)lzəni]$ is not ‘infixation’ of $els+i$ and en .

Table 190

Examples from Fabra (1956:ch.4)									
	O	D	A	X					
87		Ø _i	l _j		posa	e_j all _i	[mel]	He puts {it _j /them _j } there _i for me _k	
88	me _k	►	l _j	hi ⁺ _i		e_j e _i	[məli]		
89		►	els _j	hi ⁺ _i			[məlzi]		
90	li _k	Ø _i	Ø _j		posa	el llibre _j all _i	[li]	He puts the book _j there _i for {him _k /them _k }	
91	els _k						[(ə)lz(ə)]		
92	li _k	hi _i	Ø _j		posa	el llibre _j e _i	[li]		
93	els _k						[(ə)lzi]		
94	els _k	Ø _i	<en _i >		posa	<pa _j > all _i	[(ə)lzən]	They put some/bread _j there _i on/for them _k	
95	els _k	►	n _j	hi ⁺ _i		e_j e _i	[(ə)lzəni]		
96	els _k	hi ⁺ _i	n _j				[(ə)lzin]		
97	Qui	<els _j >	<en _j >		dóna	<pa _j > <als nois _i >?	[(ə)lzən]	Who gives some/bread _j to the them/children _i ?	
98		els _k	►	n _j	hi ⁺ _i	e_j e _i	[(ə)lzəni]	...some _j there _i on/for them _k ?	
99		els _i	Ø _j		diré		[(ə)lz(ə)]	I will tell the truth...to them	
100		hi _k	Ø _j			la veritat _j e _i	[i]	...there	
101		els _i	hi _k	Ø _j			[(ə)lzi]	...(there) on them	

Finally, els_{DAT} (99) and els_{OBL} (91) may surface as $[(ə)lzi]$ as described above. Rather than treat $[i]$ as an epenthetic vowel specific to dative l -clitics (López Del Castillo 1976),¹⁶⁴ these forms may be seen as cases of *re-analysis*. Since prosodic epenthesis produces $\{els \sim 'ls \sim lse \sim else\} / \{en \sim 'n \sim ne\}$, frequency of $(e)ls/(e)n+hi$ leads to $[əlzi]/[ni]$, *not* as i -epenthesis on $(e)ls/(e)n$, but as open-syllable allomorphs of $[əlsə]/[nə]$. In dialects where DAT.PL and ACC.M.PL are syncretic, $[(ə)lzi]_{\text{DAT}}$ opposes $[(ə)lzə]_{\text{ACC}}$, just as lis_{DAT} opposes els_{ACC} in others. In terms of paradigm uniformity, lis may be seen as adding plural-marker s to DAT.SG li , and $elsi$ as adding dative-marker i (< li/hi >) to plural els . Since all cases of $[ni]$ in the studies consulted derive from $/n'hi/$, it is unclear whether $[ni]$ has been similarly re-analysed, although Fabra's warning implies that it might.

¹⁶⁴ Elsewhere, these are always $[ə]$.

Gavarró (1992) explains Catalan *l'hi* etc. by a complex arrangement of licensing empty categories, i.e. *l'hi* is really *li+Ø*. In our account, *l'* is ACC, and *hi* is DAT. It is simply that they have swapped positions. This requires no specialized rules and can be extended to all the combinatorial changes. Far from requiring complex morphological operations, clitic weights and 3-3-rules for each dialect is all that is required to model form *and* sequence of any DAT+ACC combination across dialectal space.

6.5 Occitan

This section reviews Gascon (West), Languedocian (Central), Provençal (East). Each group has a normative/literary version, but also many dialects including A/D~D/A variations. Allocation of dialects to each group varies amongst authors e.g. Narbona/Besiérs/Montpelhiér are claimed for Languedocian (Alibèrt 1976) and Maritime Provençal (Ronjat 1913).

Table 191

Provençal	Bayle 1989; de Fourvières 1986; Ronjat 1930; Vouland 1988
Niçois	Vouland 1988; Sardou 1978
Languedocian	Alibèrt 1976
Gévaudan	Camproux 1958; Alibèrt 1976; Vouland 1988
Limousin	Chabaneau 1876[1980]; Tinton 1982
Gascon	Birabent & Salles-Loustau 1989; Lespy 1880, Rohlf's 1977
Béarnais Gascon	Hourcade 1986; Lespy 1880
Auvergnat	Bonnaud 1992
Old Occitan/Provençal	Jensen 1986; Skårup 1986; Smith & Bergin 1984

A/D-order predominates in the North. Northern and many Languedocian dialects retain the *li/lor* distinction, whilst Provençal dialects are case-syncretic for different choices of *i/ié/li*. Gascon stands out due to 3.DAT/ACC syncretism and use of *ac/ne* as default accusatives.

Table 192

		ACC.SG		ACC.PL		N	DAT.SG		DAT.PL		LOC
		M	F	M	F		M	F	M	F	
Gascon		lo/u	la	los/us	las/us/les	ac/at	lo/u	la	los/us	las/us	i/li
	Béarnais Gascon	lo	la	los	las	at/ac	lo	li/i	los	los/lis/is/i	i
Provençal	Niçois	lou		lu [ly]	li	hu	li		li		li
	Maritime	lo	la	lei(s)/li		va/vo	li/i		li/i		i (li)
	Rhodanian	lou	la	ié (i)		ac/at	ié		ié		i (li)
Languedocian		lo/le ¹⁶⁵	la	los/les	las	ò(c)	li/i		lor/li		i (li)
Northern	Gévaudan	lo	la	los	las	ò(c)	li/i		lor		i
	Limousin	lo	la	los	las	o [u]/au	el/ilh/li/i		lor		i
	Avergnat	le	la	leu	la	ò(c)	lī		lhu/lī/ī		lai/ī
Old Occitan/Provençal		lo	la	los	las	o(c)	li		lor		i

Most speakers avoid *nous/vous*, but replacements vary widely e.g. *(e)ns*, *enze*, *se*, *bous*, *-bs* [-ps/-bz], *-p/b/ts* [dz]. Niçois *nous/vous* often reduce to *n'v'*. Languedocian generally has *nos/vos*, but *nos*→*se* in the East: *s'endormirem*←*nos endormirem* (Lapalma). For this reason, the tables in this section do not include 1/2.PL forms. Their behaviour follows the same patterns as their singular counterparts.

6.5.1 Development

Old Provençal followed A/D-order (i.e. heavy datives) except with equally heavy accusatives (Jensen 1986:103; Skårup 1986:86). D/A-order for 1/2 combinations (implying light 1/2.DAT) appeared in XV^c, becoming dominant during XVIII^c. D/A for 3-3 (implying light 3.DAT) appears in XVII^c. Use of *i*⁺ for *li* (like Italian *ci*_{IMP}) is attested in [OP].¹⁶⁶ Although *li/ié/i* are sometimes treated as allomorphs, particular forms are always preferred in any given context (Bayle 1989:78; de Fourvières 1986:39; Ronjat 1930:§§497-498).

In XVI^c, DAT.PL shows both *li* and *lour*. By XVII^c, Saboly (Rhodanian) employed

¹⁶⁵ For speakers using *le(s)* as the article, this also replaces the pronouns *lo(s)*.

¹⁶⁶ Brusewitz (1905:27-29) for examples of these developments.

3.DAT.SG/PL *li*. During XVIII^c, exclusive use of *li* was established in Maritime Provençal, *ié* in Rhodanian. Some dialects retain A/D-ordering for 1/2 combinations e.g. Niçois where *li* has also spread to LOC. Dialects with *li*, retain potential substitution by *i*⁺ (transparent for *i*(é) dialects),¹⁶⁷ although ACC-ellipsis is preferred where the meaning is clear. In 3-3-contexts, dialects with *li*⁺_{DAT} show A/D except with o_{ACC}/ne_{ACC} (Gévaudan: *lou li moustrarai*), whilst those with *li*⁻_{DAT} show D/A in all circumstances (Maritime: *li lou paguè*). In both cases, *i*⁺_{LOC/IMP} advances. Niçois' *li*⁻_{LOC} means that *i*⁺_{IMP} is not available, whilst *li*⁻_{LOC} never advances.

Languedocian dialects generally retain *i*_{LOC} vs. *li*_{DAT.SG/PL} (or *li*_{DAT.SG}/*lor*_{DAT.PL}). Spoken Languedocian “confuses” *i/li* (Alibèrt 1976:64). In speech, *i* often substitutes for *li/lor* in isolation and consistently for 3-3-combinations. The distinction is generally maintained in writing but sometimes used to avoid alliteration e.g. *li+la*→*la+i*.¹⁶⁸ Conversely, *li* may replace *i* in order to avoid hiatus with *preceding* vowels.

Table 193

	LOC	3.DAT.SG	3.DAT.PL	1/2.DAT	
Old Occitan/Provençal	i(e) ⁺	li ⁺	lor ⁺	me ⁺	
Languedocian Type	i(e) ⁺	li ⁺	li ⁺ /lor ⁺	me ⁻ ~ me ⁺	XV XVI XVII
Maritime Provençal	i(e) ⁺	li ⁻	li ⁻	me ⁻	XVIII
Rhodanian/Literary Provençal	i(é) ⁺	i(é) ⁺	i(é) ⁺	me ⁻	
Niçois	li ⁻	li ⁻	li ⁻	me ⁺	
		i(e) ⁺			

En⁺_{DAT} advances over light accusatives (*l'*<*en*⁺> *tiri* <*d'acqui*>) and *en*_{ABL} follows *SE*_{NOM} (*anatz-vos-en*). As indefinite accusative *en*⁺_{ACC} follows datives (*me'n dona* ‘give some to me’), preventing any heavy datives/locatives advancing, such that *en* follows in both

167 Where it is optional, replacement by *i*⁺ is not a 3-3-rule, but selection of a different construction.

168 Similarly, Niçois, *li+la* is generally avoided by ACC-ellipsis.

ACC+en⁺_{DAT} and DAT+en⁺_{ACC}. *En* is often ‘doubled’: initially *n’en vole*, or in order to avoid hiatus with *preceding* vowels: *dunatz-me-n’en*.¹⁶⁹ The same phenomenon is seen in locative combinations: *n’i’n farai/dunatz-n’i’n*, where it also serves to maintain the *li+en/l’en~i+en/i’n* distinction, which becomes obscured in dialects where *li→i*.¹⁷⁰ The combination *ne+ne* does not occur; the result would be *n’en*, already used for *ne_{ACC}* alone.¹⁷¹

Table 194

i(é) ⁺ _{LOC}						li ⁻ _{LOC}	
i ⁺ _{DAT}		li ⁻ _{DAT}		li ⁺ _{DAT}		li ⁻ _{DAT}	
ne ⁺	ne ⁻	ne ⁺	ne ⁻	ne ⁺	ne ⁻	ne ⁺	ne ⁻
la _{ACC} +i _{DAT}		li ⁻ _{DAT} +la _{ACC}		la _{ACC} +li ⁺ _{DAT}		li ⁻ _{DAT} +la _{ACC}	
la _{ACC} +i ⁺ _{IMP}		la _{ACC} +i ⁺ _{IMP}		la _{ACC} +i ⁺ _{IMP}		(l)i ⁻ _{IMP} +la _{ACC}	
la _{ACC} +i ⁺ _{LOC}	la _{ACC} +i ⁺ _{LOC}	la _{ACC} +i ⁺ _{LOC}	la _{ACC} +i ⁺ _{LOC}	la _{ACC} +i ⁺ _{LOC}	la _{ACC} +i ⁺ _{LOC}	li ⁻ _{LOC} +la _{ACC}	li ⁻ _{LOC} +la _{ACC}
la _{ACC} +en ⁺ _{GEN}		la _{ACC} +en ⁺ _{GEN}		la _{ACC} +en ⁺ _{GEN}		la _{ACC} +en ⁺ _{GEN}	
i ⁺ _{DAT} +ac ⁺ _{ACC}		li ⁻ _{DAT} +ac ⁺ _{ACC}		li ⁺ _{DAT} +o ⁺ _{ACC}		li ⁻ _{DAT} +o ⁺ _{ACC}	
i ⁺ _{DAT} +en ⁺ _{ACC}		li ⁻ _{DAT} +en ⁺ _{ACC}	li ⁻ _{DAT} +en ⁻ _{ACC}	li ⁺ _{DAT} +en ⁺ _{ACC}	li ⁻ _{DAT} +en ⁻ _{ACC}	li ⁻ _{DAT} +en ⁺ _{ACC}	
i ⁺ _{LOC} +en ⁺ _{ACC}	n ⁺ _{ACC} +i ⁺ _{LOC}	i ⁺ _{LOC} +en ⁺ _{ACC}	n ⁺ _{ACC} +i ⁺ _{LOC}	i ⁺ _{LOC} +en ⁺ _{ACC}	n ⁺ _{ACC} +i ⁺ _{LOC}	(l)i ⁻ _{LOC} +en ⁺ _{ACC}	
Rhodanian/Lit. Provençal		Maritime Provençal		Languedocian		Niçois	

In some idiolects,¹⁷² *en*’s weight has been lost (like Barceloní, §6.4.2), resulting in *n’i⁺* joining *m’i⁺/t’i⁺* etc., and the fact that *en_{GEN}* no longer advances (*en_{ABL}* is unaffected since it is OBL). This does not, however, produce *en_{GEN}*+*la/me* etc.. In fact, use in clusters, which is always limited, seems to be replaced by use of *i(é)⁺_{LOC}*, where source/destination is read from context: *lou ié tira di man*. Other than reducing the usage of *en_{GEN}* in combinations (see also Italian, §5.2.2) the change only affects this combination and can be seen as form of ‘regularisation’ of the activity of *i⁺* in regard to accusatives.¹⁷³

169 Auger (1994:33) notes that *en* is often realized as *nn* or *n’en* in several of French varieties, including Quebec French. Penello (2004) reports similar forms *nin* in Romagnol dialects.

170 *n’i’en* before a consonant is special to Literary and Rhodanian Provençal.

171 Searches failed to find *en_{GEN}*+*o/ac_{ACC}*, possibly following from ACC specificity requirements like Catalan *ho~lo* (§6.4.2).

172 This variant was already present in [OP]: *n’i=en+y*, but *li-n/l’en=lui+en* (Brusewitz 1905:31).

173 Some dialects take the reanalysis of 3.OTHER for 3.DAT one step further, replacing the labile DAT.PL *lor* with a new form including plural morpheme *-s* giving SG~PL: *i_{DAT.SG}~is_{DAT.PL}* (e.g. *que is parlo*, ‘I speak to them’), matching languages such as Spanish *le_{DAT.SG}~les_{DAT.PL}*.

This range of subtle dialect/idiolect variation has previously been impossible to capture. Feature-based analyses are inappropriate since feature-combination \leftrightarrow surface-form relationships are many-to-many mappings. Feature combinations only *select* surface-forms, their relative weights *determine* order.

6.5.2 Provençal

Whilst D/A-order for 1/2-combinations is most common, Niçois retains A/D. The distinction affects 1/2-combinations with ACC_[+SPEC] (102) but not ACC_[-SPEC] (103-104), due to relative weight. This combines with key dialect distinctions in 3.DAT/LOC discussed above. Clearly, defining dialects in terms of A/D~D/A is meaningless.

Table 195

		ac/at ⁺	lou _{SG,M} /la _{SG,F}	li(s) _{PL,M} /lèi _{PL,F}	en ⁺ _{ACC}	i(é) ⁺ _{LOC}	en ⁺ _{GEN}
DA	me ⁻	m'at	me+lou/la	me+li(s)/lèi	m'en	m'i(é)	m'en
	te ⁻	t'at	te+lou/la	te+li(s)/lèi	t'en	t'i(é)	t'en
	se ⁻	s'at	se+lou/la	se+li(s)/lèi	s'en	s'i(é)	s'en
Rhod.	i(é) ⁺ _{DAT,SG/PL}	i ⁺ +at ⁺	lou/la+i ⁺	lis/lèi+i ⁺	i(é) ⁺ n ⁺	←n'i	
	i(é) ⁺ _{LOC}						
	en ⁺ _{GEN}		lou/la+en ⁺	lis/lèi+en ⁺			
Med.	li ⁻ _{DAT,SG/PL}	li+at ⁺	li+lou/la	li+lis/lèi	li'n ⁺	←n'i	
	i ⁺ _{LOC}	i ⁺ +at ⁺	lou/la+i ⁺	lis/lèi+i ⁺	i ⁺ n ⁺		
	en ⁺ _{GEN}		lou/la+en ⁺	lis/lèi+en ⁺			
Nic.	li ⁻ _{DAT,SG/PL}	li+at ⁺	li+lou/la	li+lis/lèi	li'n ⁺	[z]/[y] inserted as necessary	
	li ⁻ _{LOC}						
	en ⁺ _{GEN}		lou/la+en ⁺	lis/lèi+en ⁺			
AD	me ⁺	m'at ⁺	lou/la+me ⁺	li(s)/lèi+me ⁺	m'en ⁺	Swapping	
	te ⁺	t'at ⁺	lou/la+te ⁺	li(s)/lèi+te ⁺	t'en ⁺		
	se ⁺	s'at ⁺	lou/la+se ⁺	li(s)/lèi+se ⁺	s'en ⁺		

For literary Provençal, Ronjat (1913:127) notes another apparent exception to A/D order with *i(é)*_{LOC}. As shown in (§5.2.1), two locatives (subject- vs. object-oriented) are available, with different meanings. In (106), the destination *dedins* is replaced by *ie*_{LOC}, and transfer of object

to *its* resting place (object-oriented) is at issue. In (105), it is the place in which the event occurs (subject-oriented, *ié_{OBL}*) which is at issue; the destination within that place being expressed by the complement. Putative D/A~A/D-order is irrelevant.

Table 196

Table 196																			
N					O	D	A	X	N			O	D	A	X	Provençal		French	
102			me ⁻	lou ⁻				►	lou ⁻	mi ⁺	dis	Il me le dit							
103			m ^{ʔ-}	at ⁺				m ^{ʔ+}	at ⁺		doune	Je me le donne							
104			m ^{ʔ-}	en ⁺				m ^{ʔ+}	en ⁺		doune	Je m'en donne							
105		ié	Ø _i	la			ié	Ø _i	la		jito dedins _i	Il l'y jette							
106			►	la	ie ⁺ _i			►	la	ie ⁺ _i	jito e _i								
D/A Dialects					A/D Dialects														

6.5.3 Languedocian

1/2-combinations are generally A/D-order, but D/A-order appears in Cevenol [CE], and for some speakers in Foissenc/Tolaran (Alibèrt 1976). *Lor* (Foissenc/Carcassés/Albigés: *lhur*, *yur*, *lus*; Gavaudanés/Cevenol: *lür*, *lüs*) is very restricted. In Foissenc, it often combines with *i* (*lur/lus i diguèt*) corresponding to Catalan *els+hi* i.e. OBL+LOC. Whilst the written language [LG] tends to preserve *li~lor* distinctions, datives commonly reduce to *i⁺* in speech [SG]. Vowels remain in hiatus, elide, or are separated by *-z-* according to context/speaker: *ba èro/o abiò/b'auras/g'abiò/u-z-èrun/gardo-zòc*. *Nos/vos* may lose *-s*: *vo'l pòrti*, *no'ls dona*.

Table 197

		o ⁺¹⁷⁴	lo(s)	la(s)	en ⁺ _{ACC}	i ⁺ _{LOC}	en ⁺ _{GEN}
AD	me ⁺	me+o ⁺	lo(s)+me ⁺	la(s)+me ⁺	me ⁺ +n ⁺	m(e)+i ⁺	m'+en ⁺
	te ⁺	te+o ⁺	lo(s)+te ⁺	la(s)+te ⁺	te ⁺ +n ⁺	t(e)+i ⁺	t'+en ⁺
	se ⁺	se+o ⁺	lo(s)+se ⁺	la(s)+se ⁺	se ⁺ +n ⁺	s(e)+i ⁺	s'+en ⁺
LG	lor ⁺	lor ⁺ +o ⁺	lo(s)+lor ⁺	la(s)+lor ⁺	lor ⁺ +n ⁺		
	li ⁺	li ⁺ +o ⁺	lo(s)+li ⁺	la(s)+li ⁺	li ⁺ +n/en ⁺		
SG	lor ⁻	(l)i ⁺ +o ⁺	lo(s)+i ⁺	la(s)+i ⁺	i ⁺ +n/en ⁺	←-n'i DAT→i ⁺	
	li ⁻	(l)i ⁺ +o ⁺	lo(s)+i ⁺	la(s)+i ⁺	i ⁺ +n/en ⁺		
i ⁺ _{LOC} ¹⁷⁵ en ⁺ _{GEN}		(l)i ⁺ +o ⁺	lo(s)+i ⁺	la(s)+i ⁺	i ⁺ +n/en ⁺	←-n'i	
			lo(s)+en ⁺	la(s)+en ⁺			
CE	lor	lor+o ⁺	lor+lo(s)	lor+la(s)	lor ⁺ +n ⁺		
	li	(l)i+o ⁺	(l)i+lo(s)	(l)i+la(s)	(l)i ⁺ +n/en ⁺		
DA	me ⁻	me+o ⁺	me+lo(s)	me+la(s)	me ⁺ +n ⁺		
	te ⁻	te+o ⁺	te+lo(s)	te+la(s)	te ⁺ +n ⁺		
	se ⁻	se+o ⁺	se+lo(s)	se+la(s)	se ⁺ +n ⁺		Swapping

Occitan varieties have a range of upper clitic-field uses, making frequent use of OBL+DAT (107-108, note Alibèrt (1976:70)'s translations), leading to frequent clitic triplets (109-110). Many cases are ambiguous between OBL and 'ethical' datives: *me/te/nos/vos/(te+me)/(te+nos)/(vos+me)*. Their placement varies: *pòrta-i-me-ne*, *pòrta-me-i-ne*, often substituting for OBL: *se Ø/me/(te me) l'en fot*; *se (te m') i'n metèt*. Whilst these add further complexity, DAT+ACC combinations are entirely transparent, when granular weight is recognised.

Table 198

	Languedocian	French
107	Te me digue	Me _{DAT} Ø _{ACC} dites pour toi
108	Prenètz-te-me	Prenez-moi _{DAT} -Ø _{ACC} pour toi
109	Te _{OBL} l'en tiro	Il te _{OBL} l'en tire
110	Vou _{OBL} lou i'a coundu	Il vous _{OBL} l'y a conduit

174 *O* (the literary recommendation) is only used in a small part of Languedoc. Many speakers add consonants to avoid hiatus; often with pre-/post-verbal vowel variation: Albigès *ga/-gò*; Foissenc *ac/-òc*.

175 *Ye* in Agenès/Carcinòl/Albigès/Roergat: *yes dise*, *ديو-يè*, *y'abiò*.

6.5.4 Gascon

The quality of *e/a* shows wide variation, partially dependent on pre- vs. post-verbal position. In many dialects, there is little auditory difference between *los~las~les*, which may be linked to 3.DAT/ACC syncretism. Couserans has 3.DAT *li/lisi* which may also act as 3.ACC.M/F i.e. syncretism is DAT→ACC, rather than ACC→DAT as in other dialects. The following is a traditional grammar description (examples from Romieu & Bianchi 2005). Many northern dialects have replaced *ac* with *lo*, with 3-3-contexts taking *i*⁺ in a range of Catalan-like paradigms, including one where all plurals surface as *les-i* [ləzi] (Miró 2007, *in press*).

Table 199

	1	2	3	4	5	6			
ACC			lo [lu]/'l'u la [la/lɔ]/l'	nos [nus] ns [(n)s]	vos [bus] v(s) [p]	los [lus]/'ls/'us las [las/lɔs/les]	Neuter	ac/at ¹⁷⁷	ac (oc) [ɔk,ɛk]
DAT	me/m'/'m	te/t'/'t	lo [lu]/'l'u	'nse	've [pe]	los [lus]/'ls/'us	Partitive	ne/n'/'n	
REFL			se/s'/'s	[se] ¹⁷⁶		se/s'/'s	Locative	i	

Ac (111) references any gender/number and ‘matches’ *tot* (112), as *ne* ‘matches’ cardinal/indefinite adjectives. *Ne* pronominalizes inanimates *de*-phrases, partitives/indefinites (113-114), and subject attributives (115). *I* represents indefinite indirect complements (116), locatives (117), and some animate referents in 3-3-contexts.

Table 200

111	Aquò, n'ac sabi pas!	That, you do not understand it!
112	Qu'ac sabem tot sus eth!	We understand it all!
113	E me'n voletz comprar?	Do you want to buy some for me?
114	Los ne cromparà	He will buy some for them.
115	Tu que'ès gran mès jo que'n soi tanben	You are bigger than I am.
116	Qu'i pensarèi	You think so/about it.
117	Prenetz-l'i	Put it there.

176 S.W. Aquitaine *mous/se*.

177 In the North-West, *ic* ([ik]) is found in both positions.

3-3-combinations are excluded due to DAT/ACC syncretism. ACC or DAT is reduced; the results following weight order. Accusatives reduce to *ac* (determinate, 118-119) or *ne* (indeterminate, 113-114).¹⁷⁸ Alternatively, \bar{t}_{LOC}^+ is used like Catalan *hi*_{IMP} (120), overlapping with standard locative usage (117). Note that *i* is often written γ .

Table 201

When verbs license their own inherent accusative, DAT may appear alone as a 3-person personal clitic (123, syncretic with the accusative) or y_{IMP} (124). When benefactives (OBL) are present, DAT is often filled with a ‘pleonastic’ locative (125). This has the effect of making the event specific by situating it in the current time frame and of distinguishing OBL (future) from DAT (current) recipients (§3.4.3). Contra Pescarini (2015), Gascon *lou+y* and *lor+y* are not compounds, but follow the same patterns as described for Catalan *elsi/elseni*

(§6.4.2-6.4.3) and similar patterns found in Languedocian and Provençal. This combination represents *lou+y+Ø_{ACC}*, (125) or where *lou* represents syncretic 3.ACC, simple swapping of heavy *y*⁺ with light *lou* (117). This occurs more frequently in Gascon since syncretism between dative and accusative lead to frequent use of *y* for 3.DAT.

Table 202

123	Et pay lou _i Ø ditz...	Dad says to him/her _i ...	Gascon
124	Dise-y-Ø	She talks to him	
125	Lous _k y Ø _i cousinabo [de _{PRT} bounos càusos _i]	I cooked good things for them _k	
126	Ghene magno do	I eat two of them (,there)	Paduan
127	Te (*ghe)ne porto do	I bring two of them to you	

A similar effect may be seen in several Northern Italian dialects, where locative and partitive clitics are said to ‘compound’ e.g. Veneto dialects, where partitives appear as *ghe+ne* (126, Benincà 1994). When a dative is present, however, it ‘replaces’ *ghe* (127). The usage is also found in *ghe+avere* to indicate actual possession in the current situation rather than generalised ownership, like Italian *averci* (§5.5.3). An analysis based on ‘pleonastic’ use of *ghe* is more appropriate than compound forms.

As illustrated, specificity/definiteness determines clitic selection, whilst their relative weight determines order. Gascon’s apparently confusing combinatorial range is, in fact, entirely transparent, *iff* weight is recognised.

6.6 Aragonese

Aragonese¹⁷⁹ is situated between Castilian, Catalan, and Occitan, forming a dialect continuum (Kuhn 2008). External influence is reflected in clitic forms and combinations.

Table 203

	1	2	3	4	5	6		
ACC			lo/la	mos~no s ¹⁸⁰	tos~bos	los/las	Neuter	en/ne/‘n/n’
DAT	me	te	li~le			lis~les ¹⁸¹	Partitive	
REFL			se			sen	Locative	i/ie/bi ¹⁸²

Bielsa [BS] has similar clitics to Spanish plus *bi/i*, displays occasional *leísmo* (Alvar 1953:287) and D/A-ordering with no 3-3-rule, although Spanish-style spurious-*se* sometimes occurs. Ribagorza [RB] has consistent A/D clustering, but like neighbouring Catalan, DAT₃ → *i(e)*⁺. Eastern dialects of Graus and Estadilla have DAT₃₋₃ → *i(e)*⁺, but D/A-ordering. Standardised Aragonese [AR], which is close to the spoken dialect of Cheso (Landa Buil 2005; Torres Oliva 2014), is predominantly A/D-ordered with a 3-3-rule ACC → *ne*_{ACC} analogous to Gascon’s use of *o/oc*. The Zaragoza dialect [ZA] lacks this rule, leading to datives advancing over light accusatives.

179 Examples from *La Gramática de la Lengua Aragonesa* (Nagore 1977, 1989), *Conchugación de prenombrs febles de l’aragonés* (Recuenco 1992), *Las combinaciones de clíticos en el cheso* (Landa Buil 2005), *El dialecto aragonés* (Alvar 1953), *Gramática de lo cheso* (Chusé & Chuan-chusé Lagraba 1987), and (Torres Oliva 2014)’s contemporary written corpus data.

180 *Mos/tos* in eastern regions, *nos/bos* in the West.

181 Some dialects show *le/les* as in Spanish, and even *los* (as is common in Catalan).

182 Generally, *i/ie* before consonants, *bi* before *h/V*. In some regions, *bi* elides: *b’ha* (*hay*).

Table 204

	en ⁺ ACC	lo _{SG.M}	la _{SG.F}	los _{PL.M}	las _{PL.F}	i ⁺ LOC	en ⁺ GEN	
D/A	te		te+lo	te+la	te+los	te+las	t ⁺ +i ⁺	te ⁺ +n ⁺
	t/bos		tos+lo	tos+la	tos+los	tos+las	t/bos ⁺ +i ⁺	t/bos ⁺ +en ⁺
	me		me+lo	me+la	me+los	me+las	m ⁺ +i ⁺	me ⁺ +n ⁺
	m/nos		mos+lo	mos+la	mos+los	mos+las	m/nos ⁺ +i ⁺	m/nos ⁺ +en ⁺
	se(n)		se+lo	se+la	se+los	se+las	s ⁺ +i ⁺	se ⁺ n/se+ne
BS	le		le+lo	le+la	le+los	le+las		
	les		les+lo	les+la	les+los	les+las		
bi ⁺ /i(e) ⁺ LOC en ⁺ GEN	(b)i ⁺ n	lo-ye	la-ye	los-ye	las-ye			
		lo+en ⁺	la+en ⁺	los+en ⁺	las+en ⁺			
RB	li ⁺		lo-ye	la-ye	los-ye	las-ye		
	lis ⁺							
ZA	li ⁺	li-ne ⁺	lo-li ⁺	la-li ⁺	los-li ⁺	las-li ⁺		
	lis ⁺	lis-ne ⁺	lo-lis ⁺	la-lis ⁺	los-lis ⁺	las-lis ⁺		
AR	li ⁺	li-ne ⁺	li-ne ⁺	li-ne ⁺	li-ne ⁺	li-ne ⁺		
	lis ⁺	lis-ne ⁺	lis-ne ⁺	lis-ne ⁺	lis-ne ⁺	lis-ne ⁺		
A/D	te ⁺	te ⁺ +n ⁺	lo+te ⁺	la+te ⁺	los+te ⁺	las+te ⁺		
	t/bos ⁺	t/bos ⁺ +en ⁺	lo+tos ⁺	la+tos ⁺	los+tos ⁺	las+tos ⁺		
	me ⁺	me ⁺ +n ⁺	lo+me ⁺	la+me ⁺	los+me ⁺	las+me ⁺		
	m/nos ⁺	m/nos ⁺ +en ⁺	lo+mos ⁺	la+mos ⁺	los+mos ⁺	las+mos ⁺		
	se(n) ⁺	se ⁺ n/se+ne	lo+se(n) ⁺	la+se(n) ⁺	los+se(n) ⁺	las+se(n) ⁺		

/

D→I⁺

/

A→NE⁺

3-3-Rules

Swapping

Clitics precede finite (128), and follow non-finite (129-132), verbs with identical forms/sequences.¹⁸³ Datives may be doubled (137), including by impersonalizing *i(e)* in some circumstances (132). In addition to functioning as direct (133) and indirect (134) objects, *en/ne* may represent indeterminate objects (like Catalan *ho*), which remain unexpressed in Spanish (130 vs. (131)'s definite reading) and instantiate inherent accusatives, converting unergatives into transitives (135). Datives are heavy, advancing over accusatives, except heavy *ne*_{ACC}. Singular (136) and plural (137) ACC₃₋₃→*ne*_{ACC} leaving DAT unaffected, and producing surface-forms identical to partitive constructions (133).

¹⁸³ Recuenco (1992) limits *sen* to non-finite forms, but Landa Buil (2005) gives counter-examples.

Table 205

	Aragonese	N	O	D	A	X		Spanish
128				⇒	lo	me ⁺	dies	Me lo diste
129	Dando-			⇒	lo	me ⁺		Dádomelo
130	¿Quies fer-				me ⁺	ne ⁺	?	¿Quieres hacer+me+ { Ø? la?
131				⇒	la	me ⁺		
132	Enseñaz-			⇒	lo _j	ie ⁺ _i	[a los fillos] _i	Enseñadlo a los hijos
133				lis	ne		dieron tres u cuatro	Les dieron tres o cuatro
134	[De X] _j ya no'			'n _j	Ø _i		fablan	[de X] _j ya no Ø _j Ø _i hablan
135	No				en		he dormis mica	
136				li	'n		amuestro	Se lo enseño
137	Da-			lis	ne		a toz	Dáselo a todos
138		me	'n				boi ta casa	Me voy a casa
139	Ya		<'n>				viengo <de allí>	
140		se	i				caleron debaxo lo cobertizo	Se cayeron debajo del cobertizo
141	No		i				beyez cosa	No veíais nada
142	No	'n	bi				ha	No hay
143	Diners	bi	'n				ha prou	Dinero hay suficiente
144	Pueden beber	sen			ne		un baso	Pueden beberse un vaso
145			tos	se			pusieron d'acuerdo	se os pusieron de acuerdo
146				⇒	la	se ⁺	probé'n la cabeza	Se la probó en la cabeza

In the upper clitic-field, *ne* is found with SE_{ANT}+motion verbs (138), and as solitary ablatives (139) where it is unavailable in Italian. Similarly, where locatives are assumed but unexpressed in most Romance languages, 'pleonastic' *i* appears (140). With perception verbs, *i* makes constructions intransitive, with interpretations of incapacity (141, similarly in Italian). Although not mentioned by Nagore (1989), Recuenco (1992) highlights *en*'s use as indeterminate subjects of intransitive verbs (*¿Bienen ninos ta iste puesto? —En bienen*). This makes both (142-143) possible. These are not swapping, but different constructions.

Table 206

147	Se nos muere	Se nos muere
148	Me se muere o mío fillo	Se me muere mi hijo
149	Lis se i cayó	Se les cayó
150	Li se'n fue	Se le fue

Torres Oliva (2014) presents several cases as ‘A/D alternations’ compared to standard (146). Whilst most are examples of SE_{NOM} (144) or OBL+SE_{DAT} (145), (147-150) require another explanation. These do not follow normative rules, nor appear in Nagore (1989) or Recuenco (1992). Landa Buil (2005)’s study of Cheso notes OBL+*se*~*se*+OBL, but only *se*. Torres (2014)’s informants did not accept **nos*+*se*/**vos*+*se*. This cannot be N/O swapping since the putative OBL in (149-150) would conflict with *i/en*. Given the limited data, we tentatively assume, these to be examples of reflexive pronoun splitting as found in Catalan (§2.2.1).

6.7 Proclisis: Conclusions

Whilst most developments discussed above point in an A/D→D/A direction, Roergat has reduced *o*’s weight forcing the neuter into its predominantly A/D system (*zou⁻ me⁺ pagaras* (=French *tu me le paieras*). Combinations of 1/2+3 (e.g. *me+lou~lou+me*) reflect experience and influence by contact with predominantly D/A French/Spanish or A/D Aragonese. Experience of Catalan enhances the view that *i⁺* is ‘special’, promoting regularisation of its interaction with accusatives, leading to *n’i*. Speakers find an equilibrium by aligning 1/2+3 pairs with either the *en⁺/ac⁺* or *i⁺* class, or aligning 3-3-pairs with each other. Such regularisation of weight across multiple dimensions is key to describing the development of Romance clitics. Without it, analyses reduce to the itemization of random (and randomly changing) collections of rules and lexicalizations.

From the above, we argue that it is meaningless to talk about A/D~D/A languages, and fruitless to use this putative dichotomy to ‘explain’ language-specific phenomena. Each language finds an equilibrium between the weights of its clitics which is learnable, but remains open to development. As weights disappear, less evidence for them exists, and the

process accelerates in the A/D→D/A direction (i.e. underlying structural order), but as shown by Roergat, it is also possible to find/create stable states which halt the process. Such events can only occur because of the granular nature of the weight phenomenon.

6.8 Enclisis

Enclisis introduces the possibility of interaction with other pronoun types. The identification of, and sequencing effects generated by, these forms is heavily debated. We argue that once prosodic effects have been removed, all sequence changes derive directly from potentially weight-bearing allomorph selection, which is a semantic/syntactic process.

6.8.1 WP Status

(151-153) illustrate clitic~weak~strong (*gli~loro~[a loro]*) pronoun distinctions which Cardinaletti & Starke (1999) attribute to hierachical structure: (Strong(Weak(Clitic))). Manzini & Savoia (2013) provide counter-examples to the judgements upon which this hierachy is based. They argue that *loro* (<ILLORUM) is a simple pronoun (like *lui*) which has retained its ability to express *oblique* relationships. We take no position on this debate, but retain the terminology for sake of convenience.

Table 207

	[CL	CL	V	WP]	Complements
151				<loro _i >	la lettera _j <*loro _i >
152	<gli _i >	Ø _j	spedisce		la lettera _j <a loro _i >
153					<a loro _i > la lettera _j che...

From our perspective, the key factor is *placement*. Weak and strong forms are positionally distinct: *a loro* (152), but not *loro* (151) may be separated from the verb, left-dislocated, *wh*-extracted etc. *A loro* is within the complement field where it may alternate with accusative

complements based on weight (152-153); *loro* is within the verb-frame (151). Similarly for enclitics (154-155), however, *loro* does not climb (156), nor force truncation of infinitives (155), unlike clitics (154). It is, therefore, not part of the clitic-field, although it follows sufficiently closely to allow optional truncation under phrasal re-syllabification. Conversely, heavy clitics producing A/D orders can climb (157) showing that heavy clitics are not WPs.

Table 208

	CL	CL		[V	CL	CL	WP]	Complements
154			deve	spedir	<glie _i >	la _j	<loro _i >	
155				spedir(e)			loro _i	la lettera _j
156	glie	la	deve	spedire				
157	<mi	ci ⁺ >	deve	portare	<mi	ci ⁺ >		

Many varieties appear to possess accusative counterparts to *loro*_{DAT}. The sequential effects of WPs are, however, limited. WP_{DAT} causes visible change in one combination for D/A-languages, whilst WP_{ACC} effects a different combination for A/D-languages. The same changes occur if the relevant pronouns are heavy clitics (157). Evidence for WP status must, therefore, come from phonological and/or stress differences, not sequence alone.

Table 209

	D	A		WP		D	A		WP
WP _{DAT}	CL _{DAT}	CL _{ACC}		e _{DAT}	→ A+D	→	CL _{ACC}	CL _{DAT}	e _{DAT}
	Ø _{DAT}	CL _{ACC}		WP _{DAT}		Ø _{DAT}	CL _{ACC}		WP _{DAT}
WP _{ACC}	CL _{DAT}	CL _{ACC}		e _{ACC}		→	CL _{ACC}	CL _{DAT}	e _{ACC}
	CL _{DAT}	Ø _{ACC}		WP _{ACC}		CL _{DAT}	Ø _{ACC}		WP _{ACC}
									→ D+A

Ordóñez & Repetti (2006) propose that post-verbal order and stress variations derive from WPs, making phonological/prosodic processes secondary issues. They note that where proclitic and enclitic differ, post-verbal forms are always ‘fuller’ implying greater structural complexity; and if both appear post-verbally, it is the ‘fuller’ version which ‘causes’ stress-displacement and should be considered a WP. According to Ordóñez & Repetti (2006), most

D/A languages use true clitics leading to no change, whilst A/D languages (158) have generalized WPs in enclisis, which is the basis for their obligatorily final-stressing with imperatives. Although correlations exist, we argue that relationships between form, sequence and stress are not reducible in this simple manner.

Table 210

158	Bálha-lo-mé	Languedocian	Dá-lo-mé	Aragonese
	Dítz-lo-mé	Gascon	Dóna-la-mé	Mallorcan

6.8.2 L-Allomorphs & Sequence

L-allomorphs are a common ‘fuller’ form which often appear in association with stress/order changes. Corsican imperatives show intra-dialect form and sequence variation (Agostini 1984:11; Giacomo-Marcellesi 1997:21). Boucher (2013) discusses two northern speakers (from Repetti & Ordóñez (2011)’s survey) selected for displaying “a consistent pattern...not seen in speakers of all dialects”, whereby proclitic *u/a/i/e* (159) consistently alternates with enclitic *lu/la/li/le* producing A/D-order reversal (160). Following Ordóñez & Repetti (2006), Boucher equates *lu/la/li/le* with WPs and the cause of alternation, although they do not affect stress.

Table 211

159	A _{ACC} li _{DAT} kompri	You buy it for him/her/them	Corsican
160	Kompra-mi _{DAT} -la _{ACC}	Buy it for me	
161	A _[-SPEC] so	I know	
162	Un la _[+SPEC] so	I don’t know it	
163	A _[-SPEC] mi sciallu	[FR] Je me la coule douce!	

By comparison, southern dialects e.g. Gallurese have inherited *lu/la/li* from Old Corsican, whilst Modern Standard Corsican has ACC.SG *u,a,l’* and ACC.PL *i,e,l’* in both positions. This points to specialisation of existing clitics, rather separate development as WPs. Indeed, use of

the generic/neuter pronoun has been generalised across Corsican dialects, including for propositions (161) in contrast to specific items (162), and as *expletive-it* in idioms (163, see §5.5.1 for Italian *la*).

Such alternations relate to referent specificity, *not* WP status, and are common across Romance. Vinzelles (Provençal) has *u* for non-specifics, but *le* when referencing objects preceding the verb (164-165, Dauzat 1927:385,560). Whilst Provençal (§6.5.2) is recognised as having distinct *u*_{NEUT}, Corsican (re-)uses *a* which happens to be identical to ACC.F.SG (like Italian *la*). Similarly, Nuori (Sardinian) systematically represses the second [l] in 3-3 (166), but not other (167-168) contexts (Pittau 1982:83). OCP avoidance of two *l*'s (unknown elsewhere in the language) might be invoked here, or a 3-3-rule which selects a 'less-specific' ACC, much as Gascon/Aragonese select *at/ne* (§6.5.4,6.6) and Provençal/Languedocian choose ellipsis (§6.5.2,6.5.3). Either way, the effect cannot be due to WP status of the *l*-forms.

Table 212

164	Dona-me-u	Donne-moi ça	Vinzelles
165	Dona-me-le	Donne-le moi	
166	li+lu/la/los/las→[li u/a/os/as]		Nuori
167	mi+lu→[mi lu]		
168	nos/bos+lu→[no/bo lu]		
169	u/a/i cámmāni	They call him/her/them	Zonza
170	cámma-lu/la/li	Call him/her/them!	
171	dá-mmi/qđi-llu/lla/li	Give it/them to me/him!	
172	t ađđu ðittu ði dá-qđi-llu	I have told you to (= ði) give it to him	
173	un lu/la/li cámmāni	They do not call him/her/them	
174	um mi/qđi llu/lla/li ðāni mikka	They don't give it/them to me/him	
175	[→ u/a/i mmi/qđi] [ðāni	They give it/them to me/him	
176	Iđu [→ a z [Ø a llawata	He has washed it for himself	
177	[si nni [Ø kōmpra ðui	He buys some two for himself	

Zonza (Corsican, Manzini & Savoia 2015) vocalic clitics (169) incur *l*-allomorphy in modal contexts,¹⁸⁴ following imperatives (170-172), infinitives with irrealis interpretation (172, cf. Wurmbrand 2014), and preceding negated finite verbs (173-174). Datives are heavy advancing over light accusatives including *l*-less forms (175-176) producing A/D-order. *Ni*_{ACC} and *l*-accusatives are heavy, thereby retaining D/A-order as enclitics (171-172) and proclitics (174, 177). In 3-3-contexts, *(l)i*→*qi*.¹⁸⁵ Gemination may affect all consonantal pro-/enclitics with no effect on stress. Thus, *l*~*ll* alternations are determined by prosodification; *u*~*lu*~*llu* does not indicate WP status.

6.8.3 *L*-Allomorphs & Displacement

Pomaretto (Occitan) has pre- and post-verbal *l*-object clitics. Unlike proclitics (178), enclitics of all types are stressable (179). SCLs show the same alternation in stress and *l*-allomorphy (180-181). Similarly Forni di Sopra (Friulian), where *l*-less 3.NOM clitics in declarative sentences (182) alternate with *l*-forms in interrogatives (183), i.e. *l*-allomorphs are triggered by the non-veridical context of questions.¹⁸⁶ In Olivetta S. Michele (Ligurian, bordering Provençal), heavy datives advance over light vocalic accusatives producing A/D-order in both positions (184-185). However, *l*-forms (phonetically [ɾ]) can appear post-verbally inducing D/A-order (186, (Ronjat 1930). Similarly, Viozene (Imperia, Liguria) (187, Repetti & Ordóñez 2011). Classifying *rí/ré/rá/rú* as WPs might explain order change, but not post-verbal final-stress, even when putative WPs are absent (185). *l*-allomorphy (or possibly WP status) and stress are *distinct* properties.

¹⁸⁴ Other *l*-alternations are phonologically conditioned e.g. vocalic SCLs become *l* before vocalic onsets.

¹⁸⁵ Contra Pescarini (§6.3.4), this dialects shows 3-3-suppletion in A/D-order.

¹⁸⁶ Manzini & Savoia (2005:§3.6.2) for examples for numerous dialects.

Table 213

178	Lu/la/li/la: mandu	I call him/her/them	Pomaretto
179	Mandɔ-lú/ló/li/lá:/mé/nón yŋ	Call him/her/them/them/me/one of them!	
180	I(l)a: dɔrmə	They _{M/F} sleep	
181	Dɔrmən-lí:/lá:	Do they _{M/F} sleep?	
182	Al/a/i/as du'arm	S/he sleeps/They sleep	Forni di Sopra
183	Du'arm-ilu/ila/iu/ilas	Is s/he sleeping?/Are they sleeping?	
184	El u/a/i/e i/mə duna	He gives it/them to him/me	Olivetta
185	Duna-u/a/i-mé/jí	Give it/them to me/him!	
186	Duna-i-rí/ré/rá/rú	Give it/them to him!	
187	Da-rú~da-u-mé	Give it to me!	Viozene

As with interrogatives, imperatives may select particular allomorphs. In Agliano (Lucca, Tuscany), 3.M.SG proclitic $l \rightarrow lə$ (188, feminine $la \rightarrow qa$) following infinitives (193) and 1.PL imperatives (192), but $qə$ with 2.SG/PL imperatives (190-191). As (189) shows, this is not phonologically induced. Prosodically, post-verbal patterns all require a bi-moraic foot, followed by a single syllable (note the gemination in (193) to ensure this). Despite three separate forms, there is no stress-displacement. Similarly, in Anzi (194-197, Potenza, Basilicata), in addition to the $i \sim lə$ selection, 3.ACC changes with 2.SG imperatives in order to preserve its prosodic pattern. The $lə \rightarrow dde$ change may be seen as a prosodic gemination or as a requirement of WPs with 2.SG imperatives, however, its putative WP status does not induce stress-displacement. At the very most, the fixed stress-pattern influences CL~WP selection.

Table 214

188	(Nu) l vɔʝə vedé	I (do not) want to see him	Agliano
189	Lórə la cámanə	They call her	
190	(Cámə)-qə/qa	Call _{2.SG} him/her!	
191	Ca(mátə)-qə	Call _{2.PL} him!	
192	Ca(mján)-lə	Let's call _{1.PL} him!	
193	Ca(má-l)lə	To call him	
194	/vənn(i:)+ddə/	Sell _{2.SG} them!	Anzi
195	/vənn(é:)+lə/	Sell _{2.PL} them!	
196	/vənn(i:)mə+lə/	Let's sell _{1.PL} them!	
197	/nonn i vennə/	Do not sell _{2pl} them!	

Although patterns of, and triggers for, *l*-allomorphy are varied, it is clearly a product of semantic/syntactic selection between allomorphs. These ‘fuller’ forms are subject to the same prosodic effects such as gemination and change in stress (or lack thereof) as other enclitics. ‘Fuller’ forms (even in triplets, Zonza *u~lu~llu*, Agliano *l~lə~də*, Anzi *i~tə~dde*) do not *necessarily* imply WP status, and in many cases they cannot be so. In cases of possible WPs, stress is not guaranteed to change. In order to understand this variation, it is (contra Ordóñez & Repetti 2006) necessary to separate out form, sequence and stress as separate properties/processes.

6.8.4 Prosodic Structure

Apparent stress-displacement with enclitics has been addressed at length: e.g. Loporcaro 2000; Monachesi 1996; Nespor & Vogel 1986; Ordóñez & Repetti 2006; Peperkamp 1996, 1997; Torres-Tamarit 2010).

Peperkamp (1996) derives surface variation from different prosodic structures (198). This approach, however, leaves out the fact that these dialects share Three-Syllable-Rules¹⁸⁷ at PW level, providing no means to express the Three-Syllable-Rule for Lucanian, whilst Neapolitan’s inner and outer PWs have different stress rules, making penultimate stress impossible. Vogel (2009) proposes a single structure (200) where the Three-Syllable-Rule (along with segmental rules e.g. intervocalic s-voicing) is a PW property; a definition shared by all dialects, including all PWs within a dialect e.g. compounding as well as imperatives. Overall surface stress differences are expressed at CG level,¹⁸⁸ which as a distinct member of

187 Some forms allow 4th syllable stress (*péttinano*) making Peperkamp’s proposal even more problematic.

188 This prosodic constituent is *not* the ‘clitic group’ proposed by Nespor & Vogel (1986). Here, CG stands for ‘Compound Group’, representing a prosodic structure intermediate between PW and PPh.

the prosodic hierarchy has its own rules, explaining why dialects only differ as to their stress rules in the presence of clitics. Italian, which shows no stress effects, assigns CG stress to the PW's primary stressed syllable, thereby 'passing up' existing stresses. Peperkamp (1997) arrives at similar conclusions regarding the non-structured nature of clitic-fields, but retains recursion (199). However, (Loporcaro 2000:140) points out that, from Old Neapolitan (XIV^c) until the last century, attraction of stress by two clitics was not categorical for oxytone hosts e.g. ['dam:əla]~[da'm:əla]; a variant unavailable to recursive models.

Table 215

	Standard Italian	Neapolitan	Lucanian	
198	[[V] _{PW} CL CL] _{pph}	[[V] _{PW} CL CL] _{PW}	[V CL CL] _{PW}	Peperkamp 1996
199	[[V] _{PW} CL CL] _{PW}	[[V] _{PW} CL CL] _{PW}	[[V] _{PW} CL CL] _{PW}	Peperkamp 1997
200	[[V] _{PW} CL CL] _{CG}	[[V] _{PW} CL CL] _{CG}	[[V] _{PW} CL CL] _{CG}	Vogel 2009
201	vénde	cónta	vínnə	t/sell
202	véndi lo	cóntə lə	vənní llə	t/sell it
203	véndi (me lo	cóntə (mí lə	vənnə (mí llə	t/sell me it
	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;"> [[[V]_{PW} (CL CL)]_{CG} </div> <div style="display: flex; align-items: center;"> { <div> <div>Lucanian σ→[+stress]/ __syll]_{CG}</div> <div>Neapolitan σ→[+stress]/]_{PW} __syll]_{CG}</div> <div>Standard Italian σ→[+stress]/ ...]_{PW} ...]_{CG}</div> </div> </div> </div>			

Monachesi (1996) proposes that single clitics adjoin to the host forming a single PW, while clusters form a separate PW,¹⁸⁹ however, segmental rules such as intervocalic s-voicing which apply internally (204) but not across words (205), are not present in any verb/clitic combinations (206-209). The relationship which Monachesi seeks to instantiate is better expressed in terms of clusters forming independent feet, with unification of singletons (i.e. extra-metrical units) with the verb's PW taking place at the level of phrasal re-syllabification.

189 For Catalan, Torres-Tamarit (2010) propose that even clusters are part of verbal PWs.

Table 216			
204	casina	ca[z]ina	Small house
205	uovo#sodo	uovo [s]odo	Boiled egg
206	presentando#si	presentando[s]i	Presenting oneself
207	presentando#misi	presentandomi[s]i	Presenting himself to me
208	lo#sanno	lo [s]anno	They knew it
209	mi+si#dice	mi[si] dice	One tells me

Vogel (2009)’s approach allows a single analysis within a dialect for compounding and imperatives and across dialects, maintaining common features. Prosodic structure matches syntactic structure, and the concept of syntactic units (verb vs. clitic-field) is retained which is necessary given that clitics act as a group in clitic-climbing. Theory internally, it has the advantage of removing recursion.

6.8.5 Verb PW Boundary

All dialects require PWs to be at least bi-moraic,¹⁹⁰ such that not only Italian, but also Lucanian (§6.8.6), which does not possess *raddoppiamento fonosintattico*, geminate clitics following monosyllabic imperatives. Other languages employ epenthesis (Catalan, §6.4) or vowel-lengthening (Accettura, §6.8.6). This is the case whether there is stress-change or not.

Table 217													
<div><p>F</p><p>σ σ</p><p>μ</p><p>gu a (di)</p></div>	<div><p>F</p><p>σ</p><p>μ μ</p><p>d a</p></div>												
<table><tr><th colspan="3">Italian¹⁹¹</th></tr><tr><td>[[(dá.Ø)]_{PW}</td><td>]_{CG}</td><td>→ dá...</td></tr><tr><td>[[(dá.Ø)]_{PW} lo</td><td>]_{CG}</td><td>→ dál.lo</td></tr><tr><td>[[(dá.Ø)]_{PW} (te lo)</td><td>]_{CG}</td><td>→ dát.te.lo</td></tr></table>		Italian ¹⁹¹			[[(dá.Ø)] _{PW}] _{CG}	→ dá...	[[(dá.Ø)] _{PW} lo] _{CG}	→ dál.lo	[[(dá.Ø)] _{PW} (te lo)] _{CG}	→ dát.te.lo
Italian ¹⁹¹													
[[(dá.Ø)] _{PW}] _{CG}	→ dá...											
[[(dá.Ø)] _{PW} lo] _{CG}	→ dál.lo											
[[(dá.Ø)] _{PW} (te lo)] _{CG}	→ dát.te.lo											

190 Minimum word size varies cross-linguistically. Cabré i Monet (1994) proposes moraic trochees for Catalan; Thornton (1996, 2007) syllabic trochees for Italian.

191 Not shown orthographically for palatals which are always long inter-vocalically; hence *dáglielo*/**dágglielo*.

Latin imperative *-e* was lost, producing *fac*, *dīc* > Italian *fa'*, *dī'* (Mańczak 1980:68). Other short imperatives are part of a process (still productive in some dialects, Floricic & Molinu 2003) affecting frequent polysyllabic verbs e.g. *guarda/i* (<guadare)→*gua'*. In these cases, full imperative forms must be used with clitics: *guarda-lo*/**guallo*. Both variants may be understood as containing catalectic elements (cf. Kager 1995). Presence of the mora is supported not only by gemination processes but also alternations Italian *fa'*/*fai* etc. In Catalan, the *-s* of other imperatives often spreads to these monosyllabic forms, thereby restoring their minimal word-size. Over-generalisation leads to heavy imperatives in Algher e.g. *pels* 'loose' (Floricic & Molinu 2012).

Many Sardinian varieties (Pittau 1972:18-19) introduce paragogic vowels (emboldened) following stressed monosyllabic words (210), including imperatives with clitic-clusters (212), but not singletons (211). Since verb endings form the right edge of a PW, clusters form independent feet, whilst singletons remain extra-metrical. Like other Romance languages, Sardinian undergoes phrasal re-syllabification (Cardinaletti & Repetti 2009). Clitics are conjoined to the verb, inducing paragogic insertion to maintain the existing foot (212), or unification of monosyllables to create a new foot (211). It follows that (contra Monachesi 1996) association of verb and single clitics is due to phrasal re-syllabification (explaining the lack of PW-level phonology), *not* PW formation.

Table 218

210	<i>dá</i> →(<i>dáĩ</i>) _{<i>r</i>} ~(<i>dáe</i>) _{<i>r</i>}	Give!	Sardinian
211	<i>dá</i> + <i>mi</i> →(<i>dámi</i>) _{<i>r</i>} *(<i>dáĩ</i>) _{<i>r</i>} + <i>mi</i>	Give me!	
212	<i>dá</i> +(<i>milu</i>) _{<i>r</i>} →(<i>dáĩ</i>) _{<i>r</i>} (<i>milu</i>) _{<i>r</i>}	Give it to me!	

Vimeu Picard imperatives (José & Auger 2005) employ epenthesis or gemination as appropriate. Single consonant clitics e.g. *m*_{1.SG} geminate if required to fill an empty coda slot (213), whilst underlyingly geminate pronouns e.g. *ll*_{3.SG} retain both consonants, requiring epenthesis if there is no available vowel support (214).

Table 219			
213	Acoute mé bien [a(kut)(me) bjẽ] Listen to me good	Tues mmé, si tu veux [(ty.m)(me)] Kill me if you want	Vimeu Picard
214	Donne é-llé à tin père [do(n el)(le)] Give it to your father	Dis llé [(di.l)(le)] Say it	

It is often repeated (e.g. Pescarini 2015, following Teulat 1976), that Occitan shows post-imperative (never pre-verbal) optional ‘reordering’, however, these forms represent different constructions, with OBL (215) or DAT (216). The difference can be seen in their prosodic behaviour. In (215), *me*_{OBL} is an extra-metrical singleton distinct from the D/A foot, which solely contains *lo*_{ACC}. As such, *me*_{OBL} is re-syllabified to close the imperative (215), thereby losing its epenthetic vowel, at the CG level. In (216), DAT+ACC form a foot (including D/A-swapping) separated from the imperative’s prosodic word. The OBL~DAT distinction is reflected in (subtly) distinct meanings.

Table 220		
215	Daussa=m+(lo)!	Leave it for me!
216	Daussa=(lo+me)!	Leave it to me!
217	Dejá-me-(lo)!	Leave it for me!
218	Dejá-(me-lo)!	Leave it to me!

This should be compared with Spanish, which lacking D/A-swapping, has identical forms but retains the two meanings (217-218). When phrasal re-syllabification occurs in (217), *me+lo*

are run together creating the same surface form, unlike Occitan (215), where the first clitic has already been adjoined to the imperative and is, therefore, unavailable to form such a foot. The essential distinction is between extra-metrical singletons and footed pairs when CG-rules apply.

6.8.6 Lucanian

Lucanian CG's always show penultimate stress (e.g. nominal [nóčə]~[nučéd:ə] 'nut~hazel nut'), also producing stress-displacement in imperatives regardless of base stress and clitic count (Lüdke 1979). Systematic vowel changes¹⁹² as found across the language indicate that this is stress-displacement, however, it is unclear whether WPs are involved. There are informative dialect differences.

Table 221

219	u/a/lə cə:mə	I call him/her/them	Accettura
220	m u ðə:jə	He gives it to me	
221	nə mm u da	Do not give it to me!	
222	sə vənə:tə a vəder-lə	I came to see him/her	
223	[[ca:(mɛ́:) _{PW} mə)] _{CG}	Call me!	
224	[[ca:(mɛ́:) _{PW} lə)] _{CG}	Call him/her/them!	
225	[[cama:(mɛ́:) _{PW} lə)] _{CG}	Let us call _{1.PL} him/her	
226	[[cama:(tɛ́:) _{PW} lə)] _{CG}	Call _{2.PL} him/her!	
227	[[dana] _{PW} (mɛ́:lə) _{CG}	Give me it!	
228	u/a/i ɣəatsə	I lift it/them	Terranova
229	ɔ llu vi:ɣə	I don't see him	
230	ɔ mm/nɛ-u ðəvðə _{INFINITIVE}	Do not give it to me/to him!	
231	ɣwardá-llə	Look at him/her!	
232	dəna-mmíllə	Give me it!	

Accettura (Manzini & Savoia 2015) displays post-verbal /l/-allomorphy for infinitives (222) and imperatives (223-227), but not pre-verbal negators (221). Imperatives show stress-displacement, with the accent appearing verb-final, where it is not otherwise found, regardless of person (223-226), or on the cluster (227), producing the same CG-final prosodic pattern. In

¹⁹² Post-tonic vowels and pre-tonic [i,e] neutralize to [ə], pre-tonic /o/ raises, while /a/ is unaffected.

Terranova, vocalic clitics precede lexical verbs (228). High-positioned negators activate *l*-allomorphy pre-verbally (229), but not in clusters (230). Terranova has similar post-verbal stress allomorphies to Accettura for singletons (231), and clusters (232). Manzini & Savoia (2015) assume that *i:lə/illə* surfaces in (232), but elides its initial vowel in (231) in order to preserve the verb's final vowel. But the same stress pattern is induced by *l*-allomorphy in Accettura, such that Terranovan *-llə* might be the result of prosodification rather than WP status as Manzini & Savoia (2015) assume: i.e. Terranova geminates, but Accettura lengthens vowels.

Table 222

233	t-u fátstsə	I do it for you _[-SPEC]	Lucanian
234	l-ŭ fátstsə vedé	I show him it _[+SPEC]	
235	fā-m(mŭ	Do _{SG} -for me-it _[-SPEC]	
236	da-m(millə	Give _{SG} -to me-it _[+SPEC]	
237	vənnə(tillə	Sell-you it	
238	vən(nillə	Sell it	
239	mannatə(millə	Send it to me	

In the dialect presented by Ordóñez & Repetti (2006), they assume that *-illə* (236-239) is a WP corresponding to proclitic *u* (233-234), however, enclitic *u~illə* which post-verbally is determined by object specificity (235-236) might represent *u~lə*, where *lə* has geminated under stress. Indeed, Ordóñez & Repetti (2006) mention a nearby dialect of Calvello with *u~lə~illə*, which might be like Zonza *u~lu~llu* (not WP) or Anzi *i~lə~dde* (possible WP). If *illə*-forms are WPs, they sit at CG's right edge and undergo CG rules (here, penultimate stress). It does not follow that *illə_{WP}* causes stress-displacement (even less that it introduces stress); it merely provides material to which CG rules are applied. Indeed, *loro_{WP}* does not induce stress-displacement, because Italian has no such CG rule, regardless of the extra material and word-level stress made available by it.

6.8.7 Neapolitan

Whilst Neapolitan has post-imperative *l*-allomorphy with singletons (241-243), clusters seem to require ‘extended forms’ unavailable pre-verbally. Bafile (1993, 1994) assumes that *lə/la/nə* have disyllabic allomorphs *illə/ella/ennə* in clusters replacing the first clitic’s vowel, the quality of which is determined by CL₂’s gender (240-241)¹⁹³ in contrast to Lucanian’s indeterminate vowel which may be epenthetic. The implication is that, unlike Zonza *u~(l)lu*, Neapolitan has *u~lu~illu_{WP}*.

Table 223

240	dá	[[dá.m] _{PW} mə] _{CG}	[[dá.m] _{PW} (míl.lə)] _{CG}	Neapolitan
241	fǎ	[[fǎ.l] _{PW} la] _{CG}	[[fǎ.t] _{PW} (tél.lə)] _{CG}	
242	cóntə	[[cónta] _{PW} lə] _{CG}	[[cónta] _{PW} (tíl.lə)] _{CG}	
243	péttənə	[[péttəná] _{PW} lə] _{CG} → [(péttə)(nálə)] _{CG}	[[péttina] _{PW} (tíl.lə)] _{CG}	

Unlike Lucanian, imperative stress is not reduced to secondary stress as evidenced by vowels e.g. [pórta] ‘she brings’ vs. [purtátə] ‘you_{2.PL} bring’. Clusters form strong feet; singletons remain extra-metrical. Lacking Lucanian’s penultimate-stress rule, Neapolitan passes up existing stresses. PPh re-syllabification respects existing feet, but runs extra-metrical data together. If sufficient material is available (243, with proparoxytonic imperatives), new feet are created, *(péttə)(nálə)*, preserving verb-final vowel quality, without inducing gemination. Lucanian displaces stress, Neapolitan adds additional stressable positions.¹⁹⁴ There is no need to stipulate that clusters ‘select’ WPs (clitics give the same results), and no evidence that WPs effect stress patterns.

193 This is a common phenomenon. In Guardiaregia (Molise, Manzini & Savoia 2005), stressed vowels undergo metaphony, producing *i*-MASC~*e*-FEM, e.g. *da-ttʃ-iʎʎə/élla/iʎʎə/éllə*, ‘Give it/them to him!’, patterning like full pronominals e.g. *kuʎʎə/kella/kiʎʎə/kellə*. Old Neapolitan distinguished M.SG from M.PL by lack of metaphony, producing alternations such as *-mello/-millo* (Ledgeway 2009:306).

194 The intonational effects of this is discussed below.

6.8.8 Sardinian

Kim & Repetti (2013) suggest that cases in Sardinian similar to Neapolitan represent changes, not in word-level stress, but in the PPh's intonational contour, interpreted as a bitonal HL* pitch accent (also Manzini & Savoia 2005:491-505). Word-level stress remains *in situ*, usually associated with the leading tone, whilst the falling tune is associated with the rightmost metrically prominent syllable.¹⁹⁵

Sardinian has a Three-Syllable-Rule, but most words are paroxytonic. Even final stress is often converted to penultimate by adding 'paragogic' vowels /i/~e/ or /u/~o/: Campidanese *kissá*→*kissái* 'maybe' (Bolognesi 1998:66), Nuorese *kissáe*~*kissái* (Pittau 1972:19). Similarly, copy vowels are inserted after consonants in phrase-final position: *komporamídaza*~*komporamíduzu* 'buy them_{FEM}/them_{MASC} for me' (Bolognesi 1998:46). Post-verbal stress patterns vary across Sardinia. In most Logudorese/Nuorese varieties, stress remains unchanged with single enclitics (Pittau 1972:82–83; Blasco Ferrer 1988:112; Jones 1993:367). In Campidanese, placement varies with individual clitics. Clusters induce stress change in all varieties: Nuorese, Jones (1993:28); Logudorese, Blasco Ferrer (1986:114); Campidanese, Blasco Ferrer (1986:111). However, Wagner (1941:23-25) reports no such changes in Macomer (*náramilu*) and Désulo (*náramiddu*), but two accents in Campidanese *nára*+*mí*. Pittau (1972:20-21) reports both variations with proparoxytonic Nuorese verbs: *bókina*~*lu*~*bokiná*+*lu* 'call him', *bókina*+*milu*~*bokina*+*milu* 'call him for me'. Clearly, such impressionistic data requires verification.

¹⁹⁵ Prieto *et al.* (2005) for LH* pitch accent analyses of Central Catalan, Neapolitan, and Pisan.

Kim & Repetti's detailed phonetic study of Oristano (Campidanese, bordering on Logudorese) demonstrates an HL* intonational pattern. In (244-246), H associates with the verb, with L placed somewhere approaching the end of the penultimate phrase vowel,¹⁹⁶ whether clitic (244) or verb (245). Final paragogic vowels are not counted in metrical calculations (246). Their addition results in phonetic compression of syllables following V_L e.g. *d* (246) is longer than singletons, but shorter than geminates. Antepenultimate V_L (246) is shorter than penultimate V_L (245, accommodating the paragogic vowel), but still considerably longer than V_H. No compensation takes place before V_L except with monosyllabic verbs (244).

Table 224

244	[dzá _H (i) (m)mi _L ddu]	Give it to me!	Oristano
245	[abá _H ðia _L (m)mi]	Look at me!	
246	[kó _H mpora mi _L (d)dozo]	Buy them for me!	
247	[pé _M sa tí _H nde zú _L bitu]	Get up right away!	

Crucially, (247, M H+L*) shows that tones are associated with phrase-penultimate *stressable* elements, whatever word is there i.e. these are not clitic-specific patterns. Clusters, as independent feet, provide suitable anchoring points for L (244, 246) or H (247). Phrase-level re-syllabification joins extra-metrical singleton clitics (245) where verb-final vowels are elongated, acting as L's anchor. The effect is that *a* stress falls on CG's penultimate position (244-246), whether on the verb (245) or initial clitic (244, 246, 247). As indicated by vowel quality and M/H association, the original stress also remains on the verb. As long as there is sufficient distance between the two stresses, they co-exist. Monosyllabic imperatives are extended (paragogic vowel, vowel lengthening, gemination) to ensure this.

¹⁹⁶ Due to limitations of speech mechanics, tone and segment are often imperfectly aligned (Ladd 1996).

6.8.9 Sardinian II

Following Ordóñez & Repetti (2006), Hagedorn (2009) analyses Seneghese (also Oristano province) *du/doz* as WPs endowed with moraic onsets, a diachronic ‘residue’ <Latin –LL-.¹⁹⁷ Gemination is lexically-induced *raddoppiamento fonosintattico* where proclitic environments delete, whilst enclitic environments preserve and fill, the extra mora.

A simpler analysis sees gemination as prosodically induced by newly formed stressed feet. With no clitics (248), stress remains as defined by the verbal paradigm. Two clitics form an independent foot, leaving the verb’s PW unaffected (249). A singleton clitic (extra-metrical) adjoins the verb during phrasal resyllabification (250), causing changes in verb-final footing as revealed by vowel change. The same resyllabification inducing foot formation occurs for penultimate-stressed imperatives (251, note phrase-final paragogic *u*), whilst mono-syllabic imperatives are extended by paragogic *i* (252) or geminating following consonants (253) to guarantee suitable intra-stress distance. Thus, only du_{CL} is required, lengthened by post-verbal prosody, but not pre-verbally where any foot it occurs in will be unstressed relative to the phrase head i.e. the following verb.

Table 225

248	$[[\text{péttina}]_{\text{PW}}]_{\text{CG}}$	$\rightarrow [\text{péttina}]_{\text{CG}}$	Brush!	Seneghese
249	$[[\text{péttina}]_{\text{PW}} (\text{mi du})]_{\text{CG}}$	$\rightarrow [\text{péttina} (\text{míddu})]_{\text{CG}}$	Brush it for me!	
250	$[[\text{péttina}]_{\text{PW}} \text{du}]_{\text{CG}}$	$\rightarrow [\text{pétte} (\text{náddu})]_{\text{CG}}$	Brush it!	
251	$[[\text{béndi}]_{\text{PW}} \text{doz}]_{\text{CG}}$	$\rightarrow [\text{bén} (\text{díddoz}u)]_{\text{CG}}$	Sell them!	
252	$[[\text{dzá}+\emptyset]_{\text{PW}} (\text{si du})]_{\text{CG}}$	$\rightarrow [(\text{dzái}) (\text{síddu})]_{\text{CG}}$	Give it to him!	
253	$[[\text{dzá}+\emptyset]_{\text{PW}} (\text{mi du})]_{\text{CG}}$	$\rightarrow [(\text{dzám}) (\text{míddu})]_{\text{CG}}$	Give it to me!	
254	$[[\text{teléfona}]_{\text{PW}}]_{\text{CG}}$	$\rightarrow [[\text{teléfona}]_{\text{CG}}$	Telephone!	
255	$[[\text{teléfona}]_{\text{PW}} \text{mi}]_{\text{CG}}$	$\rightarrow [[\text{teléfo}(\text{ná mmi})]_{\text{CG}}$	Telephone me!	

Hagedorn (2009) mentions another local dialect Cabrarese, which shows similar accent patterns, but without post-verbal gemination. A dialect distinction based on minor

¹⁹⁷ Historically, Latin *ll* > /dd/ in Sardinia, Sicilia, and Corsica (Ferrer 1984:20).

prosodification variation (i.e. Cabrarese speakers elongate stressed vowels in preference to following consonants) seems more appropriate than one requiring distinct historical developments. (244-246) clearly show that the distinction is gradient anyway. Moreover, (253-255) shows identical patterns for *mi*, which had no means to accrue such a mora. There appears to be no reason to assume qdu_{WP} , much less that its WP status is implicated in stress-displacement.

Contra Ordóñez & Repetti (2006), apparent ‘stress-displacement’ (which never actually occurs) is not caused by qdu ’s WP status (which may be independently true), but by consistent application of prosodic rules. This helps explain the contradictory impressionistic evidence. The reporters experienced the relative prominence of two interdependent stressable positions in HL* pattern, which given different speakers and distances between stresses might be perceived as static, displaced or doubled.

6.8.10 Catalan

Catalan¹⁹⁸ uses epenthesis to ‘correct’ prosodic conditions. For Central Catalan, Campmany (2008:374) derives epenthesis (256-257) from language-wide avoidance of inappropriate intra-consonant sonority clines. Imperatives, however, require further examination. Epenthesis is also required with verbs extended by [éʃ] (Italian *-isc-*) where the resulting combination is otherwise grammatical (258), and with vocalic clitics which would normally be re-syllabified with preceding consonants, but instead become themselves syllabic (259). Moreover, the same consonant sequence may appear with/out epenthesis in different contexts (260-262).

¹⁹⁸ Examples from Bonet & Torres-Tamarit (2010).

Table 226

256	/kúz#m/	→*[kúz.mə]/ [^] [kú.zəm]		
257	/kúz#la/	→*[kúz.lə]/ [^] [kú.zələ]		
258	/sərβéz#mə/	→*[sərβéz.mə]/ [^] [sərβézəm]		
259	/kúz#u/	→*[kú.zu]/ [^] [kú.zəw]		
260	/tém-la/	→*[tém.lə]/ [^] [té.mələ]	[tém---] _{PW} lə	→[té.mə.lə] _{CG}
261	/temém-la/	→ ^v [temém.lə]/*[temémələ]	[temém] _{PW} lə	→[tə.mém.lə] _{CG}
262	/donəm#lzi/	→*[dunémłzi]/ [^] [dunéməlzi]	[duném] _{PW} lzi	→[du.ném.əl.zi] _{CG}
				→[du.ném.lo.z(i)] _{CG}

2.SG imperatives are often bare stems and hence consonant final. We posit an underlying form with an empty final vowel, ‘filled in’ at higher levels of prosody (260). This is not the case for non-2.SG imperatives, and hence epenthesis is disallowed (261), unless the clitic itself is too complex (262), in which case the clitic (*not* the imperative) undergoes epenthesis which may vary according to idiolect (Grimalt 2002). Such variations follow directly *iff* there is an imperative PW and re-syllabification at CG/PPh level (260-262).

Table 227

263	Central Catalan	[pɾumét]	[pɾumétəli]
	Formenterer	[pɾumət]	[pɾumətəli]
	Mallorcan	[pɾomət]	[pɾomətəli]

Unlike Central Catalan where CG ‘passes up’ the most prominent element, Formenterer and Mallorcan show ‘stress-displacement’ (263). Dialect-specific pronominal alternations found pre-verbally (§6.4) also appear in post-imperative (265), and post-infinitive (266, with assimilation of infinitive final *-r*) positions. Stress remains on PW (Central Catalan), but ‘shifts’ to penultimate (Formenterer, 264), or final (Mallorcan, 265), whatever happens to be there. In Mallorcan, heavy personal datives advance over light accusatives producing A/D-order and dative stress (265c), but not heavy accusatives, resulting in D/A-order and accusative stress (265d). There is no evidence that Mallorcan post-imperative stressed

pronouns (which include ACC *ó*, 265d) are WPs (contra Ordóñez & Repetti 2006). In Formenterer, stress falls on clitics in disyllabic pairs (264c), or verbs with single dative (264a) or accusative (264b) clitics, and mono-syllabic combinations (264d). Whilst Formenterer moves stress (like Lucanian), Mallorcan verb stress is not lost. Rather two stresses exist, the latter taking intonational (i.e. phrasal) prominence (like Neapolitan, Sardinian etc.).

Table 228

Formenterer			Mallorcan			
264	2.SG.IMP		265	2.SG.IMP	266	Infinitive
a)	kən(tə#lə)		[donə#lə]		[donəl#lə]	
b)	kən(tə#li)		[donə#li]		[donəl#li]	
c)	kəntə#(mə#lə)	me+le→me+le	[donə#lə#mə]		[donəl#lə#mə]	me ⁺ +le→le+me ⁺
d)	kəntə#(m#o)	me+ho ⁺ →m+o ⁺	[donə#m#ó]		[donəm#m#ó]	me ⁺ +ho ⁺ →m ⁺ +o ⁺

We conclude that sequence is determined by clitic-to-clitic relationships (weight) regardless of verb-relative position, epenthesis is determined by prosodic environment (e.g. [[V]_{PW} CL (CL)]_{CG}), and stress is determined by CG rule (Central/Mallorcan Catalan ‘pass-up’ vs. Formenterer penultimate stress). Each language then applies its own intonation pattern to the result, giving the impression of stress-displacement in Mallorcan.

6.9 Conclusions for Enclisis

In Central Catalan, tones may move for semantic effect (267, (Prieto *et al.* 2005:370), whilst spoken Spanish frequently stresses clitics following gerunds/imperatives (268, Mascaró & Rigau 2002:11). Neither is interpreted as ‘stress-displacement’. In Mallorcan, the predominant intonational stress is at the phrase’s right edge, leaving the verb *relatively* unstressed and laxing effects on vowels leading to *phonetic* reduction (269, Mascaró & Rigau 2002:11). It is this particular combination of phonetic properties consistently used in all imperatives which motivates proposals for special displacement rules.

Table 229						
267	Central Catalan	Dóna-l'hi a la Maria	H-L-L%	[do _H na.li _L a.la.ma _L %ri.a]	Neutral	Give it to Maria!
			L-H-L%	[do _L na.li _H a.la.ma _L %ri.a]	Exhortative	
268	Spanish	Cómetelo	H-L-L%	[kómetelò]	Neutral	Eat it up!
			H-L-H	[kòmeteló]	Emphatic	
269	Mallorcan	Canta!	H-L	[kántə]		Sing!
		Canta-m'ho!	H-L	[kə̀ntə mó]		Sing it to me!
		Canta'n!	H-L	/kánta+én/→[kə̀ntə́n]		Sing some!

Equally, Neapolitan/Sardinian require strong feet at their right edge which, for imperatives, is the clitic-field. Stress on this foot is perceived as stronger than verbal stress due to overall phrasal stress, which is also rightmost. No such effect occurs pre-verbally, since foot-stress on proclitics is perceived as weaker than that of the rightmost component of the group i.e. V_{FINITE} .

Unlike Sardinian/Neapolitan/Mallorcan's falling tonal patterns, Aragonese/Occitan/French have rising tones, making the effect even more marked, but still a matter of degree: "In Aragonese accents, particularly those south of Huesca and in the Ebro Valley, it is usual for the final syllable of an intonation unit, *even if unstressed*, to be given prominence by lengthening and a rise in pitch. This phenomenon, which *gives the impression of stress displacement*, is less noticeable in Chistabino" (Mott 2007:110, italics added). The left vs. right dominance of intonational patterns and depth of phonetic effects determines whether such variations are interpreted as displacement or separate phrasal phenomena. In French/Occitan (§6.10), the pattern of 'early' and 'late' rise is so marked as to be recognised as an *arc accentuel* (Fonagy 1979).

In each case, CG consists of verbal PW with its own stress, followed by a series of stressable foot heads and/or extra-metrical syllables. Languages may have CGs which adjust overall

stress placement (e.g. penultimate stress in Formenterer/Lucanian), or simply ‘pass up’ the most prominent projection(s). At PPh, re-syllabification takes place along with application of phrasal intonation. Tones are associated with stressable positions and may appear to induce stress-displacement. In fact, all stresses are still present, merely their relative prominence changes. True stress-displacement is a property of CGs and hence, like Lucanian, consistent across the language.

Whilst syntactic environment (e.g. imperative, pre-verbal negation) and context (e.g. 3-3-environments, definiteness/specificity) may change the allomorphs selected, clitic sequence always follows weight. WPs have no effect beyond adding material for CG rules to act upon (Lucanian *illə_{WP}*) or ‘pass up’ (Italian *loro_{WP}*). In short, order variation reduces to allomorph (\pm weight) selection. Accent is a product of CG rules acting upon already sequenced material. Stress is a product of tonal alignment to these already stressable positions.

6.10 French

French presents complex post-imperative variations. This section follows the arguments above, showing that prosodic structure and stress alignment are distinct from clitic sequence, repudiating WP analyses and describing the historical process which naturally engendered the particular and irregular range of patterns found in registers/dialects. Emphasis on separate enclitic series, to which dialects/registers assign different weights, provides an answer consistent with all other languages discussed in this chapter.

6.10.1 Prosodic Structure

French intonation includes an obligatory primary accent marking the right edge of prosodic

phrases assigned to its final full (non-schwa) syllable (Di Cristo 2000; Post 2000; *i.a.*). It induces syllabic lengthening, increased intensity, and unless utterance-final, rises in f_0 (Jun & Fougeron 2000). French and some Occitan dialects (Hualde 2003) also possess an optional secondary accent, an early-rise near the phrase beginning, which is not consistently accompanied by lengthening nor increased intensity, although onset consonants are often strengthened (Mertens *et al.* 2001). The early-rise (*l'accent d'insistance*) is a XIX^e development which, despite purist deprecation (Fonagy 1979), has become fully integrated even into formal speaking styles (Di Cristo 1999). Despite the name, early rises do not convey pragmatic contrasts comparable to stress languages. Fonagy (1979) opines that early- and late-rise together form an *arc accentuel* highlighting phrasal semantic unity. It has been shown to help resolve adjective scope ambiguities (Astésano *et al.* 2002; Astésano & Bard 2003) and aid word segmentation (Vaissière 1997; Di Cristo 2000). For imperatives, it reinforces the verb~clitic boundary.

Phonology reflects divisions between verbal PW and clitic-field. French final-ə is extra-metrical, unable to carry group-final accent, and regularly elides (*Puis-je* [pɥiʒ]). Such ə-elision is available for proclitics (270), but not enclitics (271). Vowel-initial clitics never resyllabify to join the imperative, rather boundaries are strengthened by *z*-insertion (274), as often found between clitics (273). The last full (non-schwa) syllable of content-words or imperatives enclitics gains phrasal-stress, lengthening and carrying the late-rise peak (Mertens 1993; Delais-Roussarie 1999; *i.a.*). Although *le* does not appear 'strong' like *moi/toi*, despite orthographic identity, imperative *le* (normally [lə]) is always [le], cannot be elided, and may take group-stress.

Table 230

270		271		272		273		274	
	(*)		*		*		*		*
	*		*		*		*		*
[(*) (*) (*)]		*) (*) (*)		[(*) (*) (*) (*) (*)]		*) (*) (*)		*) (*) (*)	
[ʒə lə vwa]		[Rǣ le mwa]		[a pɔʁt le mwa la]		[dɔn lɥi ǣ]		[dɔn zi ǣ]	
[ʒlə vwa]		[Rǣ] [mwa]				[dɔn lɥi zǣ]			
[ʒəl vwa]									
Je le vois		Rends le moi!		Apporte- le- moi là!		Donne- lui- en!		Donne- zi- en!	

Early-rises are rarely realized across proclitic function words,¹⁹⁹ but may occur in metalinguistic negation (Fonagy 1979); on determiners in enumerations/lists; and is common in television newscasters' style (Vaissière 1983). Stress does not change vowel quality, merely its length/intensity. Crucially, early-rises are more common (Delais-Roussarie 1995) on certain monosyllabic pronouns (e.g. *moi/lui*) and negative adverbs (e.g. *pas*), matching the 'special' elements of imperative contexts: *le(s)/moi/toi/lui/pas*.

Tonal attachment phonetically strengthens syllables, but does not induce changes i.e. *[lə]→[le]/*[mə]→[mwa]. Content-word schwas are not strengthened, nor receive stress. Enclitic *le* is [le] with/out stress whilst proclitic and articles are [lə] with/out stress. Moreover, *le/moi*'s realization does not change with stress placement (271, 272). We conclude that there are separate proclitic vs. enclitic series; the latter containing stressable elements. French has no CG-rule, so no accent changes are expected. Independent tonal structure simply aligns to whatever inherently stressable items are present. The remaining question is purely one of inter-enclitic sequence.

6.10.2 Against WPs

In Ordóñez & Repetti (2006)'s proposals, *moi/toi*'s WP-status explains word-order and stressability. Ordóñez & Repetti (2006) are forced to explain *moi-le* as V pied-piping WP on

¹⁹⁹ The low starting points of early-rises begin consistently at function~content word boundaries (Welby 2003).

its way to Comp, but without this highly theory-bound proposal, we are left with WPs within the clitic-field. Furthermore, *lui+le* (counterpart to *moi+le*) does not occur in most dialects/registers, requiring an unexplained distinction, even though they show identical behaviour i.e. *le+moi/lui*. Whilst *moi/lui* might each represent two different structures with their own positions and behaviour, this would have little explanatory power, providing no link between the cases.

Moreover, *le* is itself problematic; *moi/lui* cannot be separated from *le* on the basis of ‘fuller’ form or stress-ability, yet they behave differently. In (275), *le* is stressed and has a ‘fuller’ form ([*le*] not [*lə*]). If this proved WP status, (278b) is illogical; it should pattern with (277b). Moreover, order between WPs would be free (a)~(b), negating Ordóñez & Repetti (2006)’s central tenet. If *le* is a clitic (hence light, à la Ordóñez & Repetti 2006), (281a-b) follow from proclitic usage, but (279b) has no justification; it should pattern with (280b).

Table 231

275	√Régarde-le/*Régarde-lə/*Régarde'l	
276	√Régarde-moi-ça!	
	a)	b)
277	√Donne-le _{WP} -moi _{WP}	√Donne-moi _{WP} -le _{WP}
278	√Donne-le _{WP} -lui _{WP}	*Donne-lui _{WP} -le _{WP}
279	√Donne-le _{CL} -moi _{WP}	√Donne-moi _{WP} -le _{CL}
280	√Donne-le _{CL} -lui _{WP}	*Donne-lui _{WP} -le _{CL}
281	√Donne-me _{CL} -le _{LC}	*Donne-le _{CL} -me _{CL}
282	*Donne-le-MOI, pas (à) lui!	*Give it to me, not to him!
283	*Donne-le seulement moi!	*Give it only to me!
284	*Donne-le-moi et lui!	*Give it to me and to him!

As Laenzlinger (1994:85) points out, these are ‘fake’ strong forms. Despite appearances, they display clitic (not strong/weak pronoun) properties according to Kayne (1975)’s criteria: they cannot bear *contrastive* stress (282), be modified (283) nor coordinated (284). In (276), *moi* is an ‘ethical’ dative indicating the speaker’s emotional viewpoint, for which function WPs are

unavailable in French (= [CF] *à moi*). Order variation occurs with *y/en*; somewhat unlikely WP candidates, particularly since their function is regularly expressed with imperatives through *là* (272), just as WP_{ACC} is normally *ça* (276). Finally, swapping occurs pre-verbally, where WPs are unavailable in most theories. WP-status is not justified by order/stress, and simply leads to inconsistent results.

6.10.3 Development

In the earliest texts, default position for object pronouns remained post-verbal, but slowly shifted to the modern arrangement. Object pronouns were WPs (Kok 1985; Foulet 1924). D/A-order for 1/2+3 appears sporadically from XIII^c, becoming consistent during XVI^c. Old French *le(s)+me > me+le(s)*, but *la+li* remains. Whilst the earliest records followed ascending rhythm, by XIII^c, rhythm had become oxytonic with only group-final syllables bearing stress. Subject pronouns, increasingly common during XII^c, became unstressed and contractions e.g. *jol* (<*jo+le*) disappeared, leaving proclitics as an phonetically independent series. During XII^c, accent intensity weakened with various surface-form consequences, including weaker rhythmic association between object pronoun and verb.

Table 232

	IX...XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
1/2+3 order	A/D								D/A
Rhythm	Ascending	Oxytonic							
ACC-ellipsis									
	Old	Middle				XVII			

285	As me, dist il .	XI ^c	(Galambos 1985:108-112)
286	Cuide moi .	XIII ^c	
287	Il leur commanderont (< <i>le+leur</i>)	XIV ^c	
288	Je lui zi donne (=lui le)	XX ^c	

Thus, during XI^c, post-verbal pronouns had been enclitic, appearing in atonic form (285), but having lost their enclitic relationship, appeared as stressed object pronouns in group-final position (286). During the period of change (particularly XIV^c), ACC-ellipsis in 3-3-combinations was common (287), producing $li+le \rightarrow li+\emptyset$, a gap surviving in many dialects. When 3.ACC was re-introduced during XVI^c, it followed datives except $le(s)+lui$. This exception to D/A-order is retained in formal Modern French, but spoken language shows levelling towards D/A post-verbally and increasingly pre-verbally (288). Saint-Etienne French shows $lui+le$ in both positions (Morin 1979).

Rhythmic explanations (Meyer-Lübke 1899; Kukenheim 1968; Wanner 1974; *i.a.*) suggest that clitic sequence derives from oxytonic accentation, requiring heavy (*lui*) to follow light (*le*) elements. Galambos (1985:114) objects that oxytonic accentation does not require sequences of increasing heaviness except group-finally. The hypothesis could only explain post-verbal swapping, leaving pre-verbal changes as products of analogy, however, at the time of the reversal, *li* remained more frequent than *lui* even in stressed position “and *li* cannot be said to have been heavier than *le* or *la*”. Both positions represent a misunderstanding of the relationship: form \neq phonetic weight.

In our model, $li^++le \rightarrow le+li^+$ regardless of verb-relative position. It is because pre-verbal $le+li^+$ matched post-imperative $le+lui^+$, that li^+ could change to lui^+ through analogy, as illustrated by their fluctuation during this period of change. Once consolidated, lui^+ is the target form independently of the process which engendered it, and therefore, open to further independent change e.g. loss of weight producing $lui+le$. Conversely, enclitic *le* was /lə/ with

obligatory elision until XIX^c (Delais-Roussarie 1999:34) which marks the arrival of *l'accent d'insistance* where *le* becomes regularly stressed in this position. Thus phonetic form can change independently of weight, just as weight can change without effect on form.

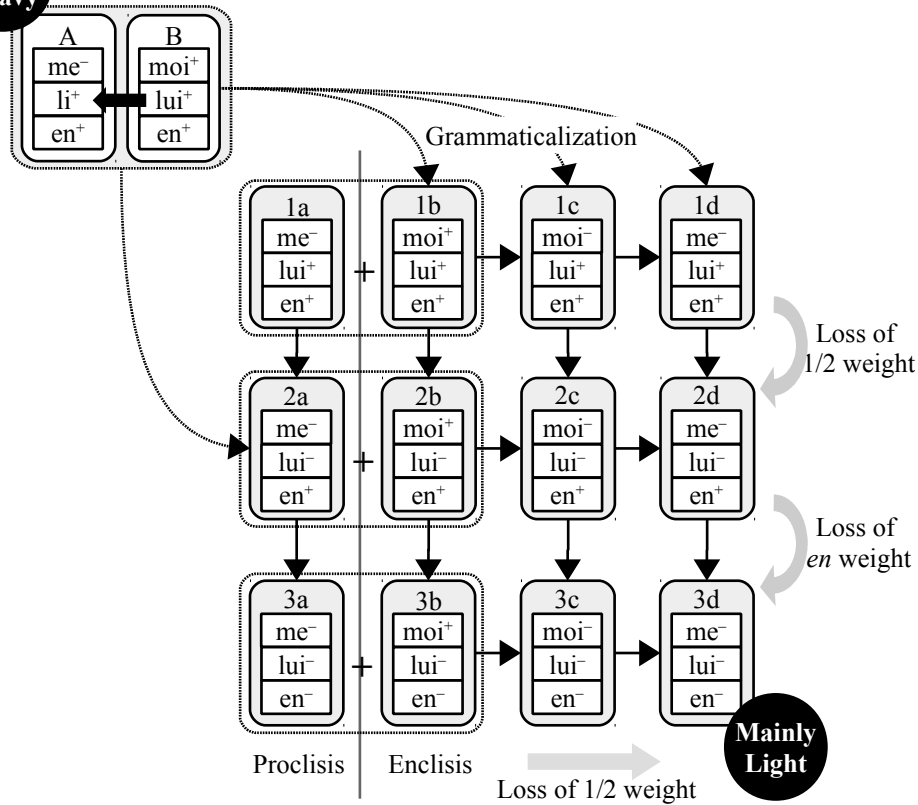
Foulet (1924) relates *me/te>moi/toi* to oxytonic phrasal stress, explaining **le>lui* as avoidance of confusion with the dative, and **le>loi* as avoidance of loss of gender distinction. The process is not, however, one of stress-induced change *e→oi*, but contextual selection. As Kayne (2000) notes, if the process were due to accentual pattern, we would expect **regarde-eux* for *regarde-les*. Formation of the enclitic series grammaticalized what was already there i.e. *moi/toi/lui* and *le* (286), not *lui* for accusative or non-existent *loi*. The pairs *regarde-les~*regarde-eux* and *regarde-le~*regarde-lui* show that enclitics are a separate series from proclitics and WPs, although they share many forms (289).

When accusatives were re-introduced to clusters, their behaviour was adapted to one of the numerous existing paradigms. The written standard 'remembered' that *li/lui* was heavy and, therefore, advanced ($\rightarrow le+li^+/lui^+$, 1a+1b). In some dialects, the accusative was simply placed in its structural position *lui+le* (2a-2b), whilst in others ACC-ellipsis *lui/leur+le→lui/leur+∅* was interpreted as a 3-3-rule. Equally, 1/2+3 were aligned to the *le+lui^+* pair ($\rightarrow le+moi^+$, 1a+1b) or structure ($\rightarrow moi+le$, 1a+1c). The latter 'supported' by apparent weight of form ([*le*], not [*lə*]). Different orders in proclisis vs. enclisis cause no conflict since they have different forms (*me~moi*), or identical form and weight (*le(s)*).

1	2	3.D	3.A	4	5	6.D	6.A	REF	GEN	LOC	NEUT
me	te	lui	lə/la	nous	vous	leur	ləs	se	en	y	le
moi	toi	lui	le/la				les	²⁰⁰	(z)en	(z)y	
						eux		soi			ça

Mainly
Heavy

Old French



290

	1a	2a	3a							
1.DAT+PRT	m'en ⁺	m'en ⁺	m'en ⁻	Pre-verbal						
GEN+1.ACC	m'en ⁺	m'en ⁺								
1.DAT+3.ACC	me+le	me+le	me+le							
LOC+1.ACC	m'y ⁺	m'y ⁺	m'y ⁺							
3.DAT+PRT	lui ⁺ +en ⁺	lui ⁻ +en ⁺	lui ⁻ +en ⁻	Shared						
LOC+3.ACC	l'y ⁺	l'y ⁺	l'y ⁺							
GEN+3.ACC	l'en ⁺	l'en ⁺								
LOC+PRT	y ⁺ +en ⁺	y ⁺ +en ⁺	en ⁻ +y ⁺							
3.DAT+3.ACC	le+lui ⁺	lui ⁻ +le	lui ⁻ +le							
	1a+1b	1a+1c	1a+1d	2a+2b	2a+2c	2a+2d	3a+3b	3a+3c	3a+3d	
1.DAT+PRT	moi ⁺ +en ⁺	moi ⁻ +en ⁺	m ⁺ +en ⁺	moi ⁺ +en ⁺	moi ⁻ +en ⁺	m ⁺ +en ⁺	moi ⁺ +en ⁻	moi ⁻ +en ⁻	m ⁺ +en ⁻	Imperative
GEN+1.ACC	en ⁺ +moi ⁺	moi ⁻ +en ⁺	m ⁺ +en ⁺	en ⁺ +moi ⁺	moi ⁻ +en ⁺	m ⁺ +en ⁺	en ⁺ +moi ⁺	moi ⁻ +en ⁻		
1.DAT+3.ACC	le+moi ⁺	moi ⁻ +le	me ⁻ +le	le+moi ⁺	moi ⁻ +le	me ⁻ +le	le+moi ⁺	moi ⁻ +le	me+le	
LOC+1.ACC	y ⁺ +moi ⁺	moi ⁻ +y ⁺	m ⁺ +y ⁺	y ⁺ +moi ⁺	moi ⁻ +y ⁺	m ⁺ +y ⁺	y ⁺ +moi ⁺	moi ⁻ +y ⁺	m ⁺ +y ⁺	
	[NF]	[CF]	[PF]	Dialect/Idiolect Variants						

200 Imperative subjects and, therefore, reflexive pronouns are *te/toi*_{2.SG}, *nous*_{1.PL}, *vous*_{2.PL}.

Similarity between proclitic $le+li^+$ and enclitic $le+lui^+$ facilitated its transposition to proclitics, however, no such path was available for $me+le/le+moi^+$, and *moi* was not accepted pre-verbally. It is also possible to ‘ignore’ the ‘emphatic’ forms like *moi* and apply proclitics in enclitic position (1a-1d), giving enclitic $me+le$ together with $le+lui/lui+le/lui+\emptyset$, according to dialect. The only impossible option is $le+me$, since it neither matches structure, nor is there any evidence for *me*’s weight in pro- or enclitic position. And this is the only pattern which does not occur.

Once weighted series were established, they continued to adapt, e.g. *lui*’s weight was lost in Quebec much later than Saint-Etienne. In some dialects, *en* became light (3a-3d, Ayres-Bennett 2004:209).²⁰¹ Type (1a+1b) represents normative style [NF], (1a+1c) colloquial usage [CF], whilst (1a+1d) is less common but also found in popular French [PF]. The remaining combinations are generally considered dialectal variations. What is most notable about (290) is not the systematic (under this analysis) differences, but rather the number of shared forms, allowing intra-dialect communication and drift. Without separate proclitic vs. enclitic series, these variations cannot be explained.

6.10.4 Analysis

In all dialects, acceptability of combinations with y^+/en^+ depends not only upon sequence (consistent with the model), but also prosodic considerations: single syllable results are heavily dis-preferred, with *-z-* often inserted. In *fausse liaison* of spoken French, post-verbal environments exhibit two *pataquès* consonants (Morin & Kaye 1982). Generally, *-t-* is associated with 3-person (291) and *-z-* with 1/2 (292). Since *z-liaison* does not generally occur

²⁰¹ The diagram shows en^+ losing weight after lui^+ as found in other languages in this chapter. I have found no evidence for the inverse order, but cannot dismiss the possibility.

in proclisis, Rooryck (1992:240-42) considers *-z-* part of imperative morphology (also Laenzlinger 1993 and Rivero 1994 for Albanian). For Laurentien French, Côté (2014) considers enclitics to have been lexicalized as underlyingly *zy/zen*. This has the unfortunate result that *y+en* (308, *-zien*) and possibly *le+lui* (319, *-zy<le+lui*) must also be treated as lexicalized pairs. Moreover, it doesn't explain cases such as *va-t'en* 'go away'. We treat *-z-* as material inserted to avoid hiatus and strengthen clitic boundaries, making *-zy/zen/zien* equivalent to Provençal *n'i/n'en/n'i'en* (§6.5.2). In formal registers, *-z-* is not recommended. Rather, vowels are elided e.g. *moi*→*m'*. Despite normative approval, many such clusters are considered unnatural and avoided.

Table 233		
291	Il devra-/t/-y avoir du monde	There must be many people
292	Donnez-le-/z/-à Marie	Give it to Mary

Combinations of *en*⁺_{GEN}+3.ACC are rare in enclisis. Morin (1979) and Grevisse & Goosse (2008) note (293)'s marginality, but acceptability with plurals (294). Similarly, combinations of *y*_{LOC}+3.ACC (295-298), although this seems to depend on verb type and/or context (299-300). A similar pattern is found in proclisis (295-296) and with 1/2.ACC. Sequences *m'y/t'y* are imposed by the norm, but generally avoided (303) by using alternative forms (305-304) or different constructions (301). Again, 1/2.ACC.PL are more acceptable (302); final-*s* of *nous/vous* (like *les*) acting as connector-*z*. LOC+PRT cluster in enclisis (307) as in proclisis (306), usually requires connector-*z* (308).²⁰² French tends to avoid *l'y*. In Old French, [li] might be *l'y* or *li*_{3.SG.DAT} (later *lui*). In Standard French, *leur*_{OBL}-*y*-Ø is found, but not **lui*_{OBL}-*y*-Ø which would be pronounced [lui]. In dialects which allow pre-verbal *z*-insertion, this often becomes *lui*_{OBL}-*zy*- Ø.

202 For Quebec French, *m/t'en*_{GEN} feel unnatural but *l'y/m'y* are simply unacceptable (Auger 1994:197-8).

In combinations of personal datives with en^{+}_{PRT} , $m/t'en$ are considered stilted in enclisis (310) and generally replaced by alternative structures (309). Because en^{+}_{ACC} is heavy, use of moi/toi always results in $moi^{\pm}+z'en^{+}$ (311-312). Similarly, for $lui^{\pm}+z'en^{+}$ (313-314), although $-z-$ does not seem to be obligatory in these cases. Since 3.ACC is light, order is determined by DAT's weight (315-318). Cases of $lui-le$ commonly appear as $z'y$, where $lui+ACC \rightarrow y+\emptyset$ (319), which may be accompanied by OBL (320, Fleurent 2015:90). The addition of moi/toi etc. to the enclitic series adds a few variations not seen pre-verbally, but where clitics match, they operate the same whatever their verb-relative position.

Table 234

	N	O	D	A	X		
293	Retire-			^(?) l'	en^{+}_{GEN}		Remove {it/them} from here
294			➡	^v les /z/			
295	Tu			^{??} l'	y^{+}	amèneras	You bring it/them there
296			➡	^v les /z/			
297	Amène-			^{??} l'	y^{+}		Bring {it/them} here
298			➡	^v les /z/			
299	Menez-			^v l'	y^{+}		Take him there
300	Conduis-		➡	^{??} les			Drive them there
301	Voulez-vous			m'	y^{+}	mener?	Will you take me there?
302	Menez		➡	^v nous /z/	y^{+}		Lead {me/us} there
303			➡	^(?) m'	y^{+}		
304			➡	moi ⁻	$z'y^{+}$		
305			$z'y^{+}$	moi ⁺			
306	Vous		y^{+}	en^{+}_{PRT}		mettrez une	You will put (some) one there!
307	Mettez- /z/		y^{+}	en^{+}_{PRT}			Put some there
308	Donne-		$z'y^{+}$	en^{+}_{PRT}			Give some there
309	Tu vas		m'	en^{+}_{PRT}		donner un autre!	You will give me another!
310	Donne-		m'	en^{+}_{PRT}			Give me some
311			moi ⁻	$z'en^{+}_{PRT}$			
312			moi ⁺	$z'en^{+}_{PRT}$			
313			lui ⁻	$(z')en^{+}_{PRT}$			Give him some
314			lui ⁺	$(z')en^{+}_{PRT}$			
315	Donne-		➡	le ⁻	moi ⁺		Give it to me
316			moi ⁻	le ⁻			
317			➡	le ⁻	lui ⁺		Give it to him
318			lui ⁻	le ⁻			
319			$z'y^{+}$	\emptyset			Give (it) there/to him/to them
320		lui	$z'y^{+}$	\emptyset			Give (it) there for him

French follows the same pattern as languages already discussed. Weight is the product of grammaticalization and may drift over time. Specific form~weight items are products of the language situation during the period of lexicalization. Change may take several paths, but all tend towards simplification (i.e. loss) of weight and hence D/A (i.e. structural) sequence, although normative stipulations often slow its advance.

6.11 3-3-Rules

This work does not attempt to ‘explain’ 3-3-rules, merely to show that once weight is taken into account, all that they represent is a set of *direct* substitutions, the range of which is more varied than previous theories allow. In addition to 3.DAT→OTHER, or no change, datives may be dropped e.g. Surmiran (321-322, Anderson 2005:243), i.e. OTHER has the surface-form Ø. As well as ‘optional’ ACC-ellipsis found in French/Provençal etc., Italo-Romance varieties such as Catanzarese (323, Pescarini 2007) and Mascioni (324-325, Abruzzi, Manzini & Savoia 2004) show systematic substitution by Ø. Alternatively, default accusatives are employed e.g. Piobbico (Marche, Manzini & Savoia 2005) which maintains 3.DAT *i*, but 3.ACC_[+DEF]→*li* regardless of ACC number/gender (326-327). See also Gascon *ac*, and Aragonese *ne*.²⁰³

Table 235			
321	Tgi dat igl matg a Gelgia?	Who is giving the bouquet to Gelgia?	Surmiran
322	Tgi igl la dat	Who it her gives?	
323	ⁿ ci+ ^l u/ndi→ ⁿ ci+Ø		Catanzarese
324	Ø lu/la/li/le a	=lo/la/li/le dà	*li+lu/la/li/le...
325	li Ø a	=gli dà/glie+lo/la/li/la dà	He gives it/them _{M/F} to him
326	m el/la/(l)i/lə 'da		He gives it/them to me
327	i li 'da		He gives it/them _{M/F} to him
328	bi/*li lu dana	nde li/*bi dana	

203 Contra Pescarini (§6.3.4), there is no relationship between A/D~D/A order and mutation.

DAT/ACC features do not solely determine their respective clitic. There is an interaction between the features of both sources upon the final output e.g. in many Catalan dialects, dative plurality determines not only dative forms but also whether accusatives are expressed (§6.4). Similarly, differences in $[\pm\text{DEF}]/[\pm\text{SPEC}]$ (329). Finally, changes in form (and thereby weight) may be accompanied by changes in order. Manzini & Savoia (2005:317-321) describe Sardinian dialects where 3.DAT $\rightarrow bi$ only with ACC $_{[\pm\text{DEF}]}$; indefinites do not trigger the 3-3-rule and li^+ advances over light *nde* (328).

Table 236

329	ACC[+SPEC]		ACC[-SPEC]	
	DAT[-SPEC]	DAT[+SPEC]	ACC[+DEF]	ACC[-DEF]
Gascon	D $\rightarrow i^+$		A $\rightarrow ac^+_{3-3}$	
Languedocian		D $\rightarrow i^+$		
Aragonese				A $\rightarrow ne^+_{3-3}$
Provençal				
Piobbico		D $\rightarrow i^-$	A $\rightarrow li_{3-3}$	
Cantanzarese		D $\rightarrow ^nci^-$		A $\rightarrow \emptyset_{3-3}$
French	D $\rightarrow y_{IMP}$		(A $\rightarrow \emptyset$)	
Italian	D $\rightarrow ci_{IMP}$	D $\rightarrow glie^-_{3-3}$		
Napoli		D $\rightarrow nce^-_{3-3}$		
Làconi			D $\rightarrow si^-_{3-3}$	D $\rightarrow ddi^+_{DAT}$

6.11.1 Putative Feature Transfer

In addition to ‘clitic fusion’ proposed for Catalan (§6.4.2), feature ‘transfer’ has been proposed for certain surface-effects found in dialects of Sardinian and Latin-American Spanish. In this section, we argue against such an analysis for Sardinian and offer a possible solution to the process as found in Latin-America. The latter proposals remain speculative since the data is too limited to make strong claims, but build on the mechanisms shown to be active in other languages, rather than adding whole new concepts to Universal Grammar just for this language.

Logudorese Sardinian is often presented as exemplifying feature-transfer i.e. the 3-3-rule generates not only 3.DAT+3.ACC→*bi*+ACC, but also DAT gender/number transfers to ACC (325, Jones 1993:220). But in many dialects *los/las* are used in isolation as datives (like Mallorcan). A more plausible analysis of (325) is as a case of ACC-ellipsis (326), following the same pattern as found with those dialects/speakers which retain *li*_{S DAT} (327). In (325-327), *bi* is OBL referencing the topic to be talked about. In true ditransitives, when accusatives are present, the 3-3-rule *li*+ACC→*bi*+ACC produces similar surface-forms (328). This is, however, coincidental as shown by dialects where *li*+ACC→*si*+ACC or *ddi*+ACC, but where (325) or (327) is still used in these circumstances. These are cases of diachronic ACC/DAT syncretism, not synchronic feature-spreading.

Table 237			
330	Nara= <i>bi</i> = <i>las/los</i> !	Tell it to them.F/M!	Logudorese Sardinian
331	Nara- <i>bi</i> - <i>los/los/la</i> +Ø _{ACC}	Speak Ø _{ACC} to them _{M/F} /her about it	
332	Narra+ <i>bi</i> + <i>lis</i> +Ø _{ACC}	= <i>dillo</i> a loro	
333	* <i>li/bi</i> l'ap <u>po</u> <u>datu</u>	= <i>gliel'</i> ho dato	

In Standard Spanish, 3.DAT+3.ACC→SE+3.ACC (the spurious-*se* rule), leaving ACC unaffected. In some Mexican/Uruguayan varieties, dative number (334) or number/gender (335-336) are said to additionally transfer to the accusative clitic (Bonet 1995a:634-635); others are restricted to number transfer (337, Kany 1951). Alonso & Henríquez Ureña (1971) include such cases in the section “Error Correction” of their grammar; Kany (1945:141) labels it a “syntactic error”; and Flórez (1977:141) states that it is “apenas pasable en el habla familiar”,²⁰⁴ but the large-scale study reported in De Mello (1992), shows that it is the most common usage in Bogota, Buenos Aires, and Mexico City.

²⁰⁴ “barely acceptable in informal style”.

Table 238

334	¿El libro _j a ellos _i quién	[a. se _i lo _j] [b. se _j los _i]	prestó?	Who lent the book _j to them _i ?
335	Si ella _i me quiere comprar el caballo _j , yo	[a. se _i lo _j] [b. se _j la _i]	venderé	If she _i wants to buy my horse _j , I will sell it _j to her _i
336	Si ellas _i me quieren comprar el caballo _j , yo	[a. se _i lo _j] [b. se _j las _i]	venderé	If they _i want to buy my horse _j , I will sell it _j to them _i
337		[a. se _i lo _j] [b. se _j los _i]		
338	Juan <se _j > <los _i > compró <un departamento _i > para <sus hijos _j >			J. bought an apartment _i for his sons _j
339	Juan <se _j > <los _i > compró <unos departamentos _i > para <su hijo _j >			J. bought some apartments _i for his son _j
340	Juan <se _j > <los _i > compró <unos departamentos _i > para <sus hijos _j >			J. bought some apartments _i for his sons _j
341	Ellos se <'lo _i /*los _i > compraron <il libro _j >			They bought the book for themselves

Oroz (1966:377) assumes that it is a response to the ambiguity caused by suppletive SE's lack of number, but in fact the process simply exchanges ambiguities. In (338-340), the number of apartments bought is no longer known (one each, or one between them). Moreover, as Company (1998:536) notes, the dative referent is always readily available in direct context, such that there is no real ambiguity to resolve.

According to RAE (1973:1571), “el plural que se observa en el complemento directo es en realidad el plural del complemento indirecto”.²⁰⁵ This, however, is not strictly true. Number is attracted to ACC only when DAT is plural. Singular datives do not overwrite the number of underlying plural ACC. Thus in (339), ACC does not reflect DAT number (i.e. it does not change to singular), and has no effect if ACC is already plural (340). Moreover, when SE is reflexive, the translation of plurality to ACC does not occur (341), adding a further restriction to the putative rule; only DAT_[-R,PL] spreads. In fact, conservation of dative plurality is a highly unlikely motivation in these dialects, where datives often do not show plurality even as singletons (§3.2.4).

²⁰⁵ “the plural observed in the direct complement is actually the plural of the indirect complement”.

Company (2001:15) considers that the “new cliticization behaves as a lexicalized, single, basically unanalyzable form: *selos, selas, seles*; in other words, *selos, selas, seles*, constitute a simplified structure, having only one object pronoun, only one argument, the Dat, which is the only pronoun that emerges morphologically, while the Acc remains inert in this grammatical area.” But if these had been lexicalized units, DAT.SG would be expected to be copied as well as DAT.PL. We do, however, agree that ACC appears “inert”.

Table 239			
	lo _{ACC}	los _{ACC}	
se ⁻ _{DAT}	se ⁻ +lo	se ⁻ +los	Standard 3.DAT[-R]+3.ACC→SE+3.ACC
le ⁻ _{DAT}	le ⁻ +lo → <i>selo</i>	le ⁻ +los → <i>selos</i>	
les ⁻ _{DAT}	le ⁻ +lo → <i>selo</i>	les ⁻ +los → <i>selos</i>	
	lo _{ACC}	los _{ACC}	
se ⁻ _{DAT}	se ⁻ +lo	se ⁻ +los	Innovative with DAT lo(s)/la(s) 3.DAT[-R]+3.ACC→3.DAT+SE
lo ⁺ _{DAT}	lo ⁺ +se ⁻ → <i>selo</i>	lo ⁺ +se ⁻ → <i>selo</i>	
los ⁺ _{DAT}	los ⁺ +se ⁻ → <i>selos</i>	los ⁺ +se ⁻ → <i>selos</i>	
	lo _{ACC}	los _{ACC}	
se ⁻ _{DAT}	se ⁻ +lo	se ⁻ +los	Innovative with DAT le(s) 3.DAT[-R]+3.ACC→3.DAT+SE
le ⁺ _{DAT}	lo ⁺ +se → <i>sele</i>	lo ⁺ +se ⁻ → <i>sele</i>	
les ⁺ _{DAT}	los ⁺ +se → <i>seles</i>	los ⁺ +se ⁻ → <i>seles</i>	

As can be seen in single clitic usage, speakers have *le/les* or innovative *lo(s)/la(s)* as dative clitics. In the proposed scenario, these clitics are considered heavy, whilst *se* remains light. The 3-3-rule has been reanalyzed replacing 3.ACC with SE (rather than 3.DAT), where SE represents a generic ACC like Gascon/Aragonese ACC→*ne/oc*, but with *se* as default accusative (aligned to its impersonal use) because of their different starting points (i.e. Spanish having lost 3.ACC.NEUT_[±DEF] centuries earlier). Because the new datives are heavy, and *se* is light (as shown by reflexive cases), they swap. The three scenarios are illustrated in (Table 239). Note only non-reflexive dative plurals ‘transfer’ their plurality. The putatively lexicalized units are exactly what they should be, and the case-model has no need for any specialised feature movement/arithmetic, fusion etc.

The data collected here, which does not represent the full range of possibilities, underscores our contention in Chapter 2 that, far from *a priori* reduction of clitics and/or classes thereof, greater sub-divisions are required. Only by comparison of *contextualised* examples illustrating the full range/combinations of feature specifications for both DAT and ACC across numerous languages and dialects (which often show subtle variations) can apparent ‘exceptions’ become part of richer patterns, and more meaningful explanations emerge.

6.12 Weight

Standard Italian *loro*_{DAT}, southern Italian dialect *ille*_{ACC} and French *ça*_{ACC} are clearly distinct from clitics in their syntactic behaviour. Whether they are WPs (Cardinaletti & Starke 1999) or simple pronouns (Manzini & Savoia 2013) is not clear. Nevertheless, we have shown that they do not affect stress, except by virtue of their length providing phonetic material which may interact with CG rules (e.g. Lucanian), but are not the source of sequence change.

Having separated out prosody from sequence, there is a need for a mechanism for changing D/A order which operates consistently pre- and post-verbally. We consider this mechanism to be clitic *weight*. Being associated with form (not features) indicates a lexicalized property, learnt by experience. It reflects distinctions present at time of creation, thereby both determining its initial range, and circumscribing future developments. That clitics can change weight whilst retaining the same form or the converse shows that these properties are distinct and should not be confused with phonetic mass or putative morphological complexity. Far from the need to postulate lexicalized pairs, it is the granularity of weight which provides for the fluidity of the overall system, and its ability to develop over time.

6.12.1 Conclusions

A single Romance-wide analysis, where languages/dialects are differentiated by their clitic lexicon including weights and separate 3-3-rules (or absence thereof) proves to be adequate to explain all variations whilst allowing us to dispense with a range of complex mechanisms which produce inconsistent results. If there is a morphological module, it only operates on 3.DAT+3.ACC which follows naturally from our model.

7 EXCLUSIONS

This chapter explores why some permutation of clitics never appears on the surface. We will conclude that there are no mechanical restrictions i.e. no syntactic/morphological mechanisms for exclusion, merely logical restrictions which take effect at the semantic level, based on the limited (i.e. disjoint vs. subset) relationships available between participants, and interlocutor perception of the linguistic and socio-linguistic situation.

7.1 Introduction

Many analyses (Bonet 1991; Grimshaw 2001; Noyer 1997; *i.a.*) treat both PCC and opacity as bans on combinations of identical or marked features. Others (Anagnostopoulou 2003, 2005; Béjar 2003; Béjar & Řezáč 2003; Ormazabal & Romero 2001, 2007; *i.a.*), derive the PCC syntactically from interaction between one agreement head and two arguments, but do not extend the analysis to clitic opacity e.g. Nevins (2007) who explicitly argues that opacity is morphological while the PCC is syntactic.

It has been shown that \emptyset (or similarly ‘generic’ forms e.g. Catalan *ho*) often surface through agreement with [–SPEC] objects (§6.4), or [–SPEC] subjects such as SE_{IMP} (§4.6.8). Chapter 6 showed that, once swapping had been removed from the picture, opacity (including \emptyset as a result) is limited to 3-3-environments. This chapter explores the remaining combinatorial effects: the PCC and putative person/number constraints. These are often overlapped, thereby obscuring simpler patterns, which we argue require no intervention by syntax or morphology.

7.2 Proposition

Bonet (1995a) defines the *strong*-PCC (*DAT+1/2.ACC) and *weak*-PCC (*3.DAT+1/2.ACC) in DOC (Double Object Constructions) constructions. Recent studies identify further variants presenting the PCC as a gradient continuum (Doliana 2013).

Table 240

DAT	ACC	Absolute	Super-Strong	Strong	Weak	Zero
1-2	3	✗	✓	✓	✓	✓
3	3	✗	✗	✓	✓	✓
1-2	1-2	✗	✗	✗	✓	✓
3	1-2	✗	✗	✗	✗	✓
Non-Romance Languages	Cairene Arabic	Kamera?				German?
Romance Languages		Surmiran?	French	Spanish	Romanian?	

Based on Kamera (Malayo-Polynesian), Haspelmath (2004) introduced *super-strong*-PCCs, which prohibit 3+3-combinations in addition to *strong*-PCC effects. However, since both objects are dative-marked (Georgi 2008 for Kamera argument encoding), and no similar cases have appeared in typological studies, its status as continuum member is problematic. We argue that Surmiran displays similar properties, however, the extra restriction is based, not on the PCC, but on 3-3-effects (§7.4.5). German, which allows any combination of objects with most verbs, may represent *zero*-PCC, however, Anagnostopoulou (2008) argues that *strong/weak*-PCC effects are present in non-default word orders. The nearest Romance case is Romanian, which we argue is *effectively* unrestricted (§7.4.4), a freedom emerging from its status as Romance's only BE-AT language. Finally, Cairene Arabic (Shlonsky 1997:207) displays *absolute*-PCC banning all cases of DOC. We know of no such case in Romance. Nevins (2007) identifies further variations including *me-first*-PCC and *te-first*-PCC, which impose local-person ordering of clitics. The existence of these variants derives from analyses of Spanish/Romanian (no other languages are given) which we believe to be erroneous (§7.3).

Restricting the analysis initially to D/A combinations, the PCC reduces to a socio-linguistic constraint upon treating [+ANIM] entities as ditransitive objects (Table 241). True breaches of this property only occur when animates are objectified. Spanish *leísta* dialects (Ormazabal & Romero 2007) demonstrate that [−ANIM] (1), but not [+ANIM] (2) objects can be transferred. In (3), animate (indicated by personal-*a*) *niña* can only be re-located. For dative clitics to appear (4), indicating reception, personal-*a* (i.e. recognition of animacy) must be removed.²⁰⁶ Inter-participant relations are expressed in terms of possession. DAT possesses ACC, which is not possible between two animate beings, except possibly infants and slaves. Similarly, inanimate objects cannot possess animate ones. In these cases, transfers can only be to/from places.

Table 241

	I	II	III _[+A]	III _[−A]	
I				✓	
II				✓	
III _[+A]				✓	
III _[−A]				✓	
LOC	✓	✓	✓	✓	

	[+ANIM] cannot possess [+ANIM]
	[−ANIM] cannot possess [+ANIM]
	[±ANIM] may be placed
	[−ANIM] objects are unrestricted

1	^v Te _{DAT} lo _{ACC+NEUT} di	[−ANIM,+recipient]	Basque Spanish
2	*Te _{DAT} le _{ACC+MASC} di	*[+ANIM,+recipient]	
3	Ø _i Ø _i llevé [a _{ANIM} la niña _{ACC}] _i [al doctor _{LOC}] _i	[+ANIM,+location]	
4	Le _j Ø _i llevé [la niña] _i [al doctor _{DAT}] _i	[−ANIM,+recipient]	

Since these restrictions hold over complements as well as clitics, non-appearance of such clitic-clusters does not reflect clitic-clitic interaction, but higher levels of syntax. In this model (contra Nevins 2007), number/person are irrelevant, since no two [+ANIM] objects can ever be combined under D/A. Putative breaches which serve to obscure this simple analysis are found with limited numbers of verbs (e.g. Spanish *presentar* ‘introduce’), but crucially, also in other non-D/A circumstances with the same apparent gaps in combinations. It is widely

²⁰⁶ Given this, [±control] may be a better description.

accepted that the PCC does not control ‘ethical/oblique’ (OBL) datives (Perlmutter 1971; Morin 1979; Albizu 1997; Ormazabal & Romero 2007; Bianchi 2006), nor nominative (NOM) clitics (Perlmutter 1971; Kayne 1975; Bonet 1991, 1995, 2008). If, as we will argue, dative forms found in these combinations are not DAT but OBL, whilst many reflexives are NOM not DAT, then such exceptions have nothing to say about the PCC.

§7.3 considers exclusions commonly attributed to number/person interaction and/or person-sequencing, showing that most may be explained by simpler and already posited mechanisms, whilst remaining cases are not conducive to such analyses. §7.4 considers apparent PCC-breaches. In our opinion, the vast PCC literature derives from erroneously mixing non-DOC cases in its description. Most putative restrictions do not hold up to empirical scrutiny, but even those that do (e.g. Spanish **me+te*) are not cases of D/A, but of N/O. It is availability of N/O and O/A constructions which allows some (but crucially, not all) combinations to carry two opposing interpretations, and engenders analyses based on person-order. Finally, §7.5.5 reconsiders whether a model of clitics requires ‘exclusions’ at all.

7.3 Person-Ordering

Non-existence of certain combinations is often offered as evidence for person-ordering e.g. Spanish **me+te~^vte+me* ‘proves’ 2»1 i.e. both underlying *me+te* and *te+me* must surface as *te+me*, thereby explaining some cases, where surface *te+me* seems to take both readings (see below). It should follow that non-syncretic **le+me~^vme+lo* ‘proves’ 1»3 and underlying *le+me* and *me+lo* must surface as *me+lo* but this is not the case: *me+lo* can never mean *le+me*. The essential logic of ‘proving’ sequences is, therefore, flawed and based on accidental syncretism.

Bonet (1995:70) notes that “there is a subset of speakers who can give [(5)] either one of two interpretations.” It is important to note that it is not stated that these speakers use such forms, rather that they can extra-linguistically generate satisfying answers to what may be unnatural questions. Unlike the *present*-type of verb discussed below, (5a/b) breaks the taboo of treating people as possessable objects which is very rare in the vast literature on the subject. A third interpretation, however, is available (6c) which matches usage when accusative complements are present (6d). (6c) uses the affectedness of *te* to imply ownership of, even identity with, the *ellipsed* accusative; ellipsis being a common means of avoiding taboo. Thus, (5b) is an inferred ‘translation’ of (6c), whilst (5a) is a literal translation of an unusual situation of overt de-humanization. The ingenuity of Bonet’s informants does not imply freedom to interpret clitics in either order (a~b) or that underlying 1+2 must surface as 2+1, as evidenced by the impossibility of reading other pairs in this fashion, or even this pair in any other situation.

Table 242

5	Te'm vendrán per divuit milions	{	a. They will sell...me to you for eighteen million	}
			b. ...you to me	
6	Te'm Ø _i vendrán el llibre _i	{	c. ...Ø _i to me on you	
			d. ...the book _i to me on you	
7	Et van recomanar a mi		They recommended {you to me/*me to you}	
8	M'ha recomanat a tu per a la feina		He has recommended me to you for the job	
9	T'ha recomanat a mi per...		...you to me	
10	%Te m'ha recomanat per...		...{me to you/*you to me}	

Acceptance of 2+1 clusters is limited. None of Martín (2012:104)’s informants did so, possibly indicating idiolect variation influenced by Spanish bilingualism. Whilst (7) can receive two interpretations for some, readings default to those with accusative clitics (8-9). If person-sequencing were active, it should be possible to read *te+me* either way, however, both authors agree that (10, Bonet 2002) has only one reading. Lack of *me+te* is, therefore, evidence of an exclusion against that combination, not enforced re-ordering.

For Spanish, Bonet takes (11-13) to imply strict person-order, but they represent different structures. Imperatives (11) are 2-person (te_{NOM}) requiring punctual events; prohibition is against achievement of a result, not against ongoing psychological processes internal to (and affecting) the listener. (11) represents a *demand* upon its recipient; only the speaker's own affectedness (me_{OBL}) matters. Conversely, statement (12) concerns that developing internal process from the speaker's (me_{MID}) perspective; listener affectedness is secondary (te_{OBL}). (13) implies the wrong relationships, i.e. it is semantically unacceptable/meaningless. It's unavailability says nothing about person-order, or even exclusions.

Table 243

		N	O	D	A	X			
11	No	te	m'				enfadis ²⁰⁷	Don't get angry on me	Punctual achievement
12	No		te	m'			enfadaré	I will not get angry on you	Inchoative process
13	*No	me	t'						Semantically incoherent

		N	O	D	A	X			
14				vi_{LOC}	$ci_{1,\text{PL}}$		manderà	He will send us to you =there/with you	
15				⇒	vi	ci^+_{LOC}		He will send you to us =here/with us)	
16				⇒	ti	ci^+_{LOC}	sei donato completamente	You devoted yourself to that/*us entirely	
17				*ti	$ci_{1,\text{PL}}$			*You gave us to yourself entirely	

		N	O	D	A	X			
18				⇒	ti	ci^+_{LOC}	presento, al direttore	I will present you to him, to the director	
19				*gli	ti				

20	[IT]	Mi [gli scrivere questa lettera]?	Would you write this letter to him for me?	1+3
21	[SP]	Se me rompió el vaso	The vase broke on me	3+1

For Italian, Bonet presents (14-15) as evidence of the *weak*-PCC, explicitly noting that locative readings are ignored. Those readings, however, are the only acceptable interpretation (16). For (14-16), ci/vi must be interpretable as a location (see translations). When this reading is unavailable (17), the sentence is ungrammatical, because its ditransitive structure is interpreted as *possessing* a person, thereby breaching $*[\pm\text{ANIM}][+\text{ANIM}]$. Matters are obscured in Italian by surface-identity of $ci_{1,\text{PL}}/vi_{2,\text{PL}} \sim ci_{\text{LOC}}/vi_{\text{LOC}}$ (unlike most Romance

207 Bonet (1991:65): these sound better in colloquial Catalan as *No se te m'enfadis* (i.e. with split reflexive).

languages), and by the fact that *ci* is heavy (§6.3), resulting in $vi+ci^+/ci^++vi \rightarrow vi+ci^+$ and $ti+ci^+/ci^++ti \rightarrow ti+ci^+$. Such cases, are not evidence for person-ordering, nor are they cases of DOC and so have nothing to say about the nature of the PCC.

Moreover, combinations banned under DAT/ACC occur in other contexts e.g. benefactive OBL (20) and anticausatives (21). Generic constraints derived from person alone are inadequate. They must be defined in terms of where they are applicable i.e. their case e.g. $gli_{DAT}+me_{ACC}$. Since the PCC already deletes $DAT_{[\pm ANIM]}+ACC_{[+ANIM]}$, none of the offending combinations can surface under DAT/ACC, so person-ordering is irrelevant and no further *X+Y style constraints are necessary.

7.3.1 Person/Number Restrictions

Other approaches treat combinatorial ‘gaps’ as complex feature processes which merely result in patterns such as *te-first-PCC* vs. *me-first-PCC* (Nevins 2007), however, empirical evidence does not support such analyses.

Table 244

Dându...	Number	Person	Mean	St.DEV.
...mi+te	SG+SG	1+2	4.91	0.36
...ți+mă	SG+SG	2+1	4.09	1.73
...mi+vă	SG+PL	1+2	3.65	1.91
...ți+ne	SG+PL	2+1	3.44	2.10
...ni+vă	PL+PL	1+2	2.97	2.12
...vi+ne	PL+PL	2+1	2.18	2.02
5-point Likert scale: 5=“completely acceptable”...1=“completely unacceptable”				

Nevins & Săvescu (2010)’s acceptability study of (non-contextualized) Romanian 1/2.SG/PL clusters following gerunds revealed significant effect for number but *not* person, *nor* number-person interaction. Whilst plurals, and particularly combinations thereof, are disfavoured, the

results are gradient; no combination is categorically (un)acceptable. It is clear, however, that 1+2 and 2+1 are equally possible, contra person-sequencing hypotheses.

Without context, listener/readers must ‘imagine’ suitable scenarios; effectively, *acceptability* becomes *likelihood of use*, not necessarily *grammaticality*. Imaginable situations might be ‘giving X to Y in marriage’, which is acceptable in Romanian, but expressed as ‘giving a hand in marriage’ in other languages. Giving children into the care of individuals/groups is also reasonable in Romanian, but not in other languages (see Spanish DOM effects, §7.2). Giving many to many is far less likely since it potentially breaches RND; if the result is seen as (i.e. it depends on listener perception of context) the union of the groups. Although differences are small, within each pair, treating *you*_[±PL] as objects is less *acceptable* (possibly considered less polite) than talking about oneself in this way. Adding the universal preference of SG>PL, provides an adequate (if not mathematically specific) analysis of the empirical continuum. The data is *not* evidence for discrete (i.e. feature-based) combinatorial restrictions on person/number.

Table 245

22

	[-R]		[+R]	
	Mean	St.Dev.	Mean	St.Dev.
Te me	1.56	1.066	5.06	2.361
Te nos	1.46	0.966	4.58	2.426
Os me	1.06	0.306	1.17	0.545
Os nos	1.00	0.000	1.65	1.037

23

[+R]		{	os+nos>os+me	NOM+OBL/DAT
[-R]	Te+me>te+nos		os+me>*os+nos	OBL+DAT
[-R]			*os+me>*os+nos	OBL+ACC

24

NOM+DAT	te+me	{	os+me<te+nos	<	os+nos
NOM+OBL			te+nos<os+me	=	os+nos

25

	Sample low rater (participant 101)				Sample high rater (participant 118)			
	Te me	Te nos	Os me	Os nos	Te me	Te nos	Os me	Os nos
OBL+ACC	1				3	7	1	
OBL+DAT					1	3		
REF+DAT					7	7		
REF+OBL	3				9	7	3	

For Ibero-Spanish, Alba de la Fuente (2012) presents acceptability tests of contextualized pre-verbal 2+1 clusters. As found across the literature, reflexives are preferred (22). Whilst there is clear bias against $os_{[\pm R]}$ as initial clitic, the only secure conclusion is $*os_{[-R]}+nos$. All other combinations are marginally acceptable to someone sometimes. The breakdown in (23) shows bias towards particular case/constructions and that reflexivity can reverse $os+nos \sim os+me$ acceptability. In fact, participant 118 also deemed non-reflexive $te+nos \geq te+me$.²⁰⁸ On-line reading tests of reflexive examples reveal processing time variations (24), whereby most commonly used $te+me$ is easiest to parse; least frequently experienced $os+nos$ is most difficult; whilst the middle variation again reflects differences in case/construction.²⁰⁹

Participants showed different response patterns (25); giving low rates to all clusters, or high rates to $te+me/nos$ (particularly reflexives), but low (or non-existent) rates to $os+me/nos$. This was unrelated to age, sex, or origin. Low-raters may have interpreted ‘acceptability’ as ‘fitting grammarian rules’, however, acceptability is clearly not determined by person/number combinations or sequencing. This data does not support any discrete ‘rules’.

Singular te/me are inherently individualised. Plurals do not denote multiples of *I* or *you* but classes from which individuals are drawn. Combinations of SG+PL, and particularly PL+PL, pronouns may be contextually read as overlapping sets (e.g. *I/We* and *You_{PL} → We together*) thereby breaching RND, whilst SG+SG requires specific context to gain such readings, usually involving different constructions (see 180, and discussion). Reflexivity guarantees

²⁰⁸ Nicol 2005 and Bianchi 2006 for similar Italian idiolect variation.

²⁰⁹ Note that, for dialects which have replaced os/nos with se , $os+me/nos$ may never have been experienced.

disjoint sets (A=SUBJ/B≠SUBJ) and, therefore, will be generally more acceptable. This effect is enhanced by proscription of OBL_[-R]+ACC vs. acceptance of NOM_[+R]+OBL.²¹⁰ All other combinations are potentially acceptable. Gradient (non-)acceptance (using *individual scales*), follows from language-wide distributional skews against plurals and social pressure, which limit experience of certain combinations and lead to their questionable status, particularly in these artificial situations. The total absence of *me+te* (not tested here) remains unexplained, but it clearly *cannot* be ‘built into’ a wider number/person-based set of rules.

7.4 Present-Verbs

Putative breaches to the *strong*-PCC are limited to a few verbs which select locations, as evidenced by their derivations and the clitics available for repairs e.g. French *y*. PLACE may be represented as locational destinations (DAT_[+E] including impersonal *y*_{IMP}) or event coincidence OBL_[+E], depending upon context and language resources. The following examples are from Italian, but identical verbs/patterns recur across Romance, and beyond: e.g. English *introduce* is unavailable in DOC constructions: *I introduce* {[√]*him to you*/**you him*}.

Talmy (1985) notes a division between Romance and Germanic with respect to lexicalization patterns:²¹¹ verb-framed languages (Romance) tend to incorporate direction/path into verbs leaving manner/instrument as adjuncts. Satellite-languages (Germanic) tend to do the opposite, leaving direction/place as adjuncts or particles (e.g. jump off): English *John danced into the room* is expressed as *John entered the room dancing* in Romance. All verbs under consideration subsume prepositions indicating incorporation of PLACE: *presentare*<prae+*sento*, *affidare*<ad+*fidare*, *raccomandare*<*re*+con+*mandare*. The roots are

210 3-3-triggers show similar ‘dependence’ on reflexives: [√]*se+lo(s)*/[√]*le(s)+se*/**le(s)+le(s)*/**le(s)+lo(s)*.

211 See Folli (2000) and Mateu (2000) for discussion and qualifications to this simple dichotomy.

activities, but additional prefixes have an “*funzione perfettivizante*” (Munaro 1994).

Table 246

26	G. ha fornito merce avariata a P.	G. has furnished damaged merchandise (to P.)
27	G. gli ha fornito merce avariata	
28	G. ha <u>ri</u> fornito Paolo di merce avariata	G. has furnished P. with damaged merchandise

In (26), the at-issue relationship is between subject and object, with no guarantee that P. has received the goods or is affected.²¹² P. is an optional destination which, when absent, is read as discourse-*here*. The clitic in (27, *gli*_[DAT, +E]) creates recipient readings; P. is the new possessor and thereby *affectee*. It is exactly these clitics (personal datives) which cannot appear with *present*-verbs. Since Italian has no *gli*_[DAT, +E], it is impossible to express P. as goal clitic.²¹³ In (28), the at-issue relationship is between G. and P. as directly *effected*, emphasized by the *ri*-prefix. (26~28) represent locative alternations similar to English *Load the hay onto the cart*~*Load the cart with hay*, whilst (27) is ditransitive.

Table 247

29	Ti Ø _i raccomando [questo libro] _i	I recommend this book...to you ([<i>all'attenzione</i> [di qn]])
30	Ti Ø _i raccomando [di non fare tardi] _i	I recommend to you not to come in late
31	Raccomandare qn a (le cure di) qn	To entrust someone[to (the care of) someone]
	Vi Ø _i raccomando [il mio bambino] _i	I entrust {my child to you/you with my child}
32	[X _i] mi Ø _i raccomando e _i !	Don't forget!
33	Mi raccomando! non perderlo	Please, don't lose it!
34	Raccomandarsi [alla piet� [di qn]]	To implore someone's pity (≈ <i>fare esortazioni</i> , plead with)
35	Mi raccomando a lei	I commend myself to you
36	Affidare un incarico a qn	To entrust somebody with a task
37	Assegnare alle cure di qn	To entrust [to (the care of) somebody]
38	Affidarsi a	(≈ <i>confidare in</i>) (place) trust in somebody/something
	Mi affido alla tua discrezione	I rely on your discretion

Raccomandare (≈*consigliare*) is generally paraphrased ‘recommend/suggest’ (29), ‘warn’ (30), or ‘entrust’ (31), where the object may be animate in restricted circumstances. A PLACE is implicit in all readings which defaults to the addressee’s memory/sensibilities (e.g. *I*

212 Affectedness, signalled by clitic-doubling in Spanish/Romanian (§3.2.5), is not allowed with these verbs.

213 We argue that Romanian does possess personal DAT_[+E] leading to its ‘freedom’ from the PCC (§7.4.4).

(re)call to your attention), but may be made overt (31,34). Reflexivity may indicate subject involvement whilst the object to be kept in mind is inferred from context (32), or added as an emphatic after-thought (33). Alternatively, reflexivity may indicate that the subject is the element to be kept in mind; (35) is a common formal salutation in letters.²¹⁴ *Affidare* follows the same pattern (36-38). Explicit cases indicate that these verbs select ACC+LOC, where LOC may be defined with reference to third-parties. Syntactically, any person subordinated to PLACE cannot be extracted to DAT_[-E] (31), but may be referenced as OBL_[+E] to the event, from which possession of the PLACE as mental location is inferred.

Table 248

39	Presentare qc in un'esposizione	Show/display something at an exhibition
40	Presentare qn in società	Introduce someone into society
41	Presentarsi davanti al tribunale	Appear before the court
42	Presentarsi a (elezione)	Stand for election
43	La situazione si presenta difficile	Things look a bit tricky
44	Presentarsi alla mente	Come/spring to mind (idea)
45	Se mi si presenterà una simile occasione	Should a similar opportunity occur/arise
46	È così che ti presenti?	Is this any way to be seen?
47	Presentarsi bene/maledettamente	Have a good/poor appearance
48	Si presentano all'improvviso	They turned up unexpectedly (=apparire, appear/turn up)
49	Presentare qn (a qn)	(≈far conoscere qn) introduce somebody
50	Presentarsi (a qn)	(≈farsi conoscere dicendo il nome) (formal) present yourself
51	Gli si è presentato come dottore	He presented himself as a doctor to/on him _{OBL}

Frequent translation of *presentare* as 'introduce' gives a false impression of reciprocity. Whilst *introduce* incorporates *into*, *present* indicates disclosure *within* a situation and can be read as bringing the object *into* view, or *to* the attention/awareness of (and hence *before*) situational attendees. By adding the prefix to its stative base, *present* comes to denote an achievement measured in terms of delivery to that PLACE. *Presentare*'s basic sense is to *make known/disclose* an object (39, ≈*esibire* [*all'attenzione* [*di qn*]]), where PLACE defaults to discourse-*here*, but may be made explicit and include people as reference points (39-42).

²¹⁴ Swiss French has similar usage: *se rappeler au souvenir de qqn*, 'recall yourself to the notice of someone'.

The secondary nature of such referents becomes clear with inanimate subjects (43-45, *≈apparire*, ‘appear’). In (45, *≈capitare*, ‘occur/arise’), *occasion* cannot give itself into *mi*’s possession (DAT): *mi* is event coincident (OBL). With animate subjects, *presentarsi* often has an intransitive quality (46-48, *≈farsi vedere*, ‘appear’), where PLACE or personal reference points are understood. Similarly for transitive cases (49-50).

Presentare is often treated as a speech verb, however in such verbs, ‘inherent’ (e.g. words) or explicit accusatives are metaphorically transferred to necessarily present recipients. With *present*-verbs, knowledge of the object (e.g. name (50) or role (51)) is simply declared. The relationship at issue is between subject and declared object, *not* optional attendees to whom there is no sense of transfer. Indeed, the object explicitly stays where and with whom it is. These are AT/WITH, *not* TO/FROM, relationships.

7.4.1 Cases

Languages may avoid PCC-clashes by leaving complements unpronominalized. Řezáč (2007, 2008) and Béjar & Řezáč (2003, 2009) *i.a.* discuss such constructions in terms of last resort phenomena used to ‘repair’ cluster violations, however, we consider complement cases as the base forms. The question becomes: what are the limitations upon conversion of complements to clitics, individually and/or in pairs? This requires an understanding of complement properties, which we argue are not as they are generally considered.

Most commonly (52), ACC is pronominalized, leaving secondary arguments as *a*-phrases (Evans *et al.* 1978:167; Seuren 1976:60). Some languages allow WPs to replace *a*-DPs, but this is generally less acceptable (Wanner 1974; Evans *et al.* 1978). Since Spanish [+ANIM] *a*-

DP_{ACC} is homophonous with *a*-DP_{DAT}, *a*-DP_{OBL}, *a*-DP_{LOC}, complement case cannot be determined from surface-form, and a second reading is marginally available (53). Similarly for WPs.

Speech-act context is central to available interpretations. Note, not only the change in functions, but also the preposition's translation. To use 'to' in English for 'with' when Carlos is present would imply the wrong relationship between presenter~presentee~audience. Thus, there are two schema against which the same surface-form may be matched with opposite meanings which must be differentiated by context. The fact that Spanish *a* is used in both circumstances merely serves to obscure the situation.

Table 249

Table 249

52	^v Te _{ACC} presento a _{OBL} Carlos	53	[?] Te _{OBL} Ø _i presento a _{ACC} Carlos _i		Spanish
	^v Ti _{ACC} Ø _i presento a _{OBL} Carlos _i		[?] Ti _{OBL} Ø _i presento Carlos _i		Italian
	[%] mi _{ACC} presentan ti _{OBL}		[?] mi _{OBL} presentan ti _{ACC}		
	I will present you to/before Carlos	}	I present Carlos to/before you	C. is absent	
	I present you with Carlos			C. is present = ‘this is Carlos’	
54	M _j ’ Ø _i ha donat [el regal _i] a mi _j		She has given [the present _i] [to me _j]		Catalan
55	<*Me _j > te _i va recomanar e _i a mi _j		He recommended...you to me		
56	En Josep _i , me’l _i va recomanar		...him to me		
57	*A en Josep _i , me li _i va recomanar		*...me to him		
58	<*Le _j > me recomendó <a él _j >		...me to him		Spanish
59	(*Le _j) <me _i > recomendó <a mi _i >		...me _i (*to him _j)		
60	J. y M. le han sido presentados (a K.)		J. and M. was introduced to K.		
61	I. și M. și-au fost prezentați		J. and M. were presented to each other		Romanian

Whilst *te* (53) might be DAT or OBL, *a Carlos* (52) cannot be DAT, since Spanish requires dative-doubling. Similarly, Catalan dative-doubling is obligatory with strong pronouns (54, Bonet 1991:204-5). Their absence with *present*-verbs (55) indicates that these are not DAT_[-E]. Because of 1/2 syncretism, *me/te*'s case is not surface apparent, but its syntactic behaviour is

not that of DAT, as shown in 3-person, where generally 3.DAT \neq 3.ACC. In the absence of reflexives (see below), Spanish *le(s)*_{DAT} is *never* (Perlmutter 1971; Bonet 1991) available (58-59). Similarly, Catalan *present*-verbs take *el*_{ACC} (56) but not *li*_{DAT} (57).²¹⁵ Thus, in the one situation where DAT is demonstrable, it is ungrammatical. These are not DAT_[-E], but OBL. As shown in §3.4.1, passives do not take DAT, only OBL. The fact that the secondary complement of *present*-verbs shows in passives (60) with the same meaning as in active forms confirms that these verbs do not select for DAT. Alone of all the Romance languages, Romanian possesses personal locative clitics i.e. DAT_[+E] (61, Dobrovie-Sorin 2006:132), allowing it to apparently breach the PCC (§7.4.4).

7.4.2 Constructions

Reflexives introduce restrictions on role interpretation, but greater numbers of potential constructions. As OBL+ACC_[+R], they follow the same pattern, but in circumstances where the reflexive cannot be ACC, the empirical generalization is: “If the linear cluster order is indirect object second, then the indirect object is [–Reflexive]” (Evans *et al.* 1978). In our terms, NOM_[+R]+OBL_[-R].

For many speakers, SE_{NOM} is available with *present*-verbs to intensify, or show subject involvement in, the event. In (62), *se* cannot refer to presentee (*lo*) or audience (*a ellos*), but only the subject-presenter (NOM_[+R]). *Ellos* cannot be recipients: logically, they do not possess *lo*; nor grammatically, since such datives must be doubled in Spanish. *Ellos* are ‘on stage’ but the ‘spot-light’ is on the relationship between subject-presenter (V_N) and object-presentee (*lo*); a monovalent process *before/in the presence* of *ellos* i.e. third parties are referenced as place (OBL_[+E]). Since Spanish does not possess 3.OBL_[+E], clitic versions are unavailable (63),

²¹⁵ The same holds for French *lui* and Italian *gli*.

unlike other persons (64) and in contrast to $^{\vee}le+se$ (OBL_[-E]) in affectee constructions. Thus $*le+me/te$ with *present*-verbs does not reflect the PCC, but an independent lack of resources, observable in all OBL_[+E] uses. That these ‘datives’ are OBL is confirmed by the impossibility of ‘present X before Y on Z’ where ‘before Y’ and ‘on Z’ would both occupy OBL position, in contrast to their availability in ditransitives where recipient (DAT) and event affectee (OBL) may co-exist. If the audience were DAT, le_{DAT} would be unrestricted.

Table 250									
		N	O	D	A				
62	SE _{NOM}	[se _i]	Ø _k	I	le _j	presentó _i e _j	a ellos _k]	He <i>presented</i> him to them
63			*les _k				e _k		
64			<me _k >				<a mí _k >		...to me
65	Neutral	[Ø _i]	Ø _k	I	lo _j	presentó _i e _j	a ellos _k]	He presented him to them
66			*les _k				e _k		
67			<me _k >				<a mí _k >		...to me
68	SE _{ACC}	[Ø _i]	<me _k >	I	se _i	presentó _i e _j	<a mí _k >]	He presented himself to me
69			<*les _k >				<a ellos _k >		
70			Ø _k						...(to current audience)
71	SE _{ANT}	[se _i]	<me _k >	I		presentó _i]	<a mí _k >]	He declared himself before me
72			<les _k >				<a ellos _k >		
73			Ø _k						...(to current audience)
74			le _k				[al juez _k]		He appeared...before the judge
75			Ø _k				[a las autoridades _k]		...before the authorities

The same logic holds for neutral transitives (65-67=62-64 without SE_{NOM}), but there is no sense of subject involvement. Alternatively, presentee and subject may be identical, referenced by SE_{ACC}, whilst third parties continue to be referenced by OBL (68-69). Again, $le(s)_{OBL}$ is unavailable, because the referent is a place (OBL_[+E]), not affectee (OBL_[-E]).

Finally, *present*-verbs may be used intransitively with SE_{ANT} (71-72), describing a subject COS potentially *affecting* on-stage third parties, which are represented by OBL_[-E].²¹⁶ This

²¹⁶ Note change of translation to indicate the internal nature of the event.

means that *le(s)*_[OBL,-E] and combinations such as *me/te/se+le(s)* become available with this particular meaning (71), overlapping with surface forms containing polite *le*_{3.ACC} with the opposite meaning (62-64). The distinction is confirmed by doubling. As discussed in Chapter 3, whilst OBL_[+E] cannot be doubled (62-69), OBL_[-E] with SE_{ANT} may depending on the type of referent: (74, affectee)~(75, unaffected destination/replicate mass).²¹⁷ Note that (71-75) cannot be ACC/DAT since they would be in reverse order, nor DAT/ACC since the meaning would be inverted.

It is immediately evident how, out of context, identical surface-forms can often represent direct and inverted relationships between participants. This is not the result of syntactic, morphological or prosodic processes, nor free interpretation. The clitics are where they should be. The listener may match the same surface-sequences to different constructions. Due to the nature of the verbs, semantic differences are limited and communication is not impaired should the listener select a construction different from speaker intentions; they represent nuances giving prominence to different participants. The greater the context, the fewer possible interpretations. It is only out of context that any ambiguity arises and acceptability judgements become a game concocting suitable scenarios to fit randomly selected sequences i.e. the activity becomes linguistically meaningless.

The possibility of multiple readings with such verbs should be compared with those of true ditransitives e.g. (76, Nicol 2005:190) which is not ‘ungrammatical’, but is semantically strange, as becomes clear when components properties are highlighted in the translations.

²¹⁷ Surface identical (70~73) differ in perspective. In (70), the subject ‘introduces himself’, rather than waiting for someone else to do so, whilst in (73), he ‘makes himself known’ to those present e.g. *De repente, Juan entró a la fiesta y comenzó a presentarse.*

Table 251

76

??Te me muestras en el espejo

Te _{NOM}	me _{OBL}	You are (going and) _{NOM} showing	yourself in the mirror	on/for me _{OBL}
Te _{NOM}	me _{DAT}			to me _{DAT}
Te _{NOM}	me _{ACC}		me in the mirror	
Te _{DAT}	me _{ACC}	You are showing me in the mirror to/for yourself		

Nevertheless, the restrictions that do exist are sufficient to produce some ‘gaps’ in available choices, noted in the literature, but for which previous models can only provide stipulations. In this analysis, systematic restrictions such as presence/absence of *le* and why only some combinations have two readings, emerge naturally.

The empirically noted preponderance of reflexives may represent a preference for less ambiguous forms. However, normative proscription against OBL+ACC_[±R] in general and OBL+ACC_[+R] in particular (§1.2.3) may counter this. In semantic terms, *me_{NOM}+te_{OBL}*/*me_{ANT}+te_{OBL}*/*te_{OBL}+me_{ACC}* *presento* differ in that SE_{NOM} implies stronger agentivity, SE_{ANT} implies a subject-oriented view, whilst *te_{OBL}+me_{ACC}* presents an object-oriented view, making each more/less appropriate to each situation. Thus, even in languages/dialects which have the capability to express the full range of constructions, acceptability remains context- and speech-act dependent.

7.4.3 Western Romance

For Catalan, otherwise ungrammatical *me+li* becomes acceptable with ‘ethical’ datives (77) and ‘inherent’ reflexives (78). (79) is marginally acceptable, particularly if subject involvement is emphasized (*jo mateix*). For (78-79), many speakers prefer *hi*. Generally, dative clitics cannot resume complements; only *hi* is used as a place/situation reference (81). We consider *li* in (79) to be OBL_[−E] in a personal anticausative construction like (78). With *hi*,

it might be read as (80), with no change in participant interaction, but a slight change in emphasis. As noted (§7.3), 2+1 acceptability is limited but where found, it follows the same pattern of OBL+ACC (82, only one reading) vs. SE_{ANT}+OBL (83, only one reading).

For Italian, Seuren (1976) only accepts reflexive structures, whilst Evans *et al.* (1978) merely sees them as favoured (86). Both authors only accept *mi+ti* (84). Seuren (2009), however, notes increased marginal use of (85). Italian, therefore, displays opposite properties to Spanish: Spanish $\sqrt{te+me} \sim {}^{\circ}me+te$ vs. Italian ${}^{\circ}ti+mi \sim \sqrt{mi+ti}$. (85) is unavailable in Standard Italian which only has *mi*_{OBL}, but acceptable in dialects/idiolects which also possess *ti*_{OBL} (§3.3.5). This reflects the clitic lexicon and need not be expressed in featural terms.

Table 252

	N	O	D	A			
77	No	me	li			diguis que calli	Don't tell him/her to shut up on me
78	A la Roser	me	li			vaig declarar	I declared myself (my love) to Rose
79	A en Pere					vaig recomanar jo mateix	I remembered myself to Pere
80			► m'	hi ⁺		vaig recomanar	I recommended <i>myself</i> to Pere
81	A la Roser		hi			parlaré demà	As for T, I will talk with her tomorrow
82		${}^{\circ}te$	m'			recomanat per a la feina	He recommended {me to you/*you to me}
83		${}^{\circ}te$	m'			vas presentar a la festa	You introduced {yourself to me/*me to you}
84		mi	ti _[-E]			raccomando	I commend/remember myself to you
85			${}^{\circ}ti$ _[+E]		mi		
86			${}^{\circ}mi$ _[+E]		ti	raccomandano	They recommend you to me

The PCC is considered absolute in all French varieties (Morin 1979; Quicoli 1982, 1984; Burston 1983) and Old French (Jensen 1986). Potential breaches (87) are expressed by accusative clitics with other parties in *à*-phrases. Since the second participant does not possess ACC, it cannot cliticize to DAT. *À*-phrases represent looser connections, which we consider to be OBL_[+E]/DAT_[+E]. As noted in (§3.3.5), French/Italian do not accept SE_{ANT}+OBL_{PERSONAL}, so this pattern is unavailable in these languages.

French lacks personal locative clitics (88), but does have y_{IMP} (89). For some e.g. colloquial Parisian French (Couquaux 1978:211-213), y_{IMP} may also replace animate participants (90).²¹⁸ Kayne (2008:182) observes that this is common across Romance as a marked option. Even with *penser*, y can have local-person referents (91-92) in restricted circumstances, becoming more available in coordination and clitic left-dislocation (92). Couquaux reports that the same speakers who reject (90) also reject (91), suggesting that y is the same in both contexts (Postal 1990 for similar arguments). In clusters (even more rare), y^+ advances over light ACC (94), showing that these are object-clitics. Whilst (94) might be SE_{NOM} , (90) can only be ACC. Seuren (1976:11) specifically notes the impossibility of 1/2.ACC in clusters including reflexives (93). Absence of such forms must be attributed to lack of suitably locative clitics ($OBL_{[+E]}/DAT_{[+E]}$), as found in other languages. This leaves y_{IMP} as the only available means of pronominalizing such participants in French; hence the absolute nature of its PCC.

Table 253

87	Il vous présentera à moi	He will introduce you to me
88	Je < *lui > pense < à lui >	I think about him
89	Je < y > pense < à cette question >	I think about this matter/it
90	Il m'y a présenté (?y=à eux / ?y=à vous)	He showed me to them/you _{PL}
91	Il y _i pense, { à elle _i /vous _i }, toutes les nuits	He thinks of her/you _{PL} every night
92	Je pense à toi _i et j'y _i penserai toujours	I think of you and I always will
93	*Il { me se/se me } présente	*He presents himself to me
94	Il s'y présente (y ⁺ _{DAT} +se _{ACC} →s'y ⁺)	He presented himself there (=before relevant person(s))

Maritime Provençal is similar to French, but because se_{ACC} is heavy, $li_{LOC}+se^+_{ACC}$, rather than $*s'i^+_{LOC}$ (95, =French *Il s'y présente*). Rohlf's (1977:182) provides examples of benefactive OBL with y in Gascon (96). Aragonese, which has a tendency to 'pleonastic' $(b)i$,²¹⁹ shows $se+i$ combinations susceptible to two context dependent readings (97-100). Other cases

218 Foulet (1919:§436) notes that i for *lui* has been attested since Old French: *Mes ge la vi e s' i parlai*, 'but I saw her and spoke to her'.

219 This may indicate that default person/place for OBL in Aragonese is i , not \emptyset .

clearly indicate $SE_{ANT+OBL}$ (101). OBL status for audiences is confirmed by (102), where DAT+ACC would induce swapping (§6.6). In Barceloní Catalan, hi_{IMP} frequently appears in such situations (103), whilst acceptance of Italian ci_{IMP} varies. Bianchi (2006: 2039) accepts (106), which Cardinaletti (2008:45) specifically rejects.

Recipient/possessive datives cannot be repaired with y/hi . Postal (1990) and Řezáč (2007) *i.a.* take this as evidence that PCC ‘repairs’ involve realization as PPs, since ‘repairable’ datives are those that alternate with a +DPs thereby excluding these datives in French. However, Catalan possessive datives may appear as full a +DPs, with/out doubling (104-105, Rigau 2002:2076). The relationship to a +DPs is, therefore, determined by the language’s dative-doubling capabilities, not repair strategies. We consider y/hi ’s inability to ‘repair’ possessives, but appearance in *present*-type clauses, as evidence that the ‘repairable’ items are not datives, but locatives. Standard French which has no $OBL_{[+E]}$, has no means, to extend monovalents present-verbs and hence never ‘breaches’ the PCC. What is seen in colloquial French/Catalan/Italian is extraction of the secondary adjunct as locative $y/ci/hi_{IMP}$ as long as the referent is easily obtained from context.

Table 254

	N	O	D	A	X		
95			li_{LOC}	se		presènto	Provençal
96	...que	les	y	\emptyset_i		presentèc [era siebo fénnou] _i	Gascon
97		\emptyset_j	⇒	se _j	i^+	presientan [propuestas concretas] _j	Aragonese
98		se _i	i	\emptyset_j			
99		\emptyset_j	⇒	se _j	i^+	sinyoron [30.519 contractos nuevos] _j	
100		se _i	i	\emptyset_j			
101	En esta ocasión	se _i	mos			presentan [fixaus y contrastaus] _i	
102	Xordica		mos	la _i		presenta e_i agora en una edizi3n m3s complleta	
103			⇒	m'	hi^+	ha recomanat la senyora	Barceloní
104	En Pere		(li _i)	\emptyset_j		dibuixa un palasso _j	[a la seva filla] _i
105						renta la cara _i	
106			⇒	ti/vi	ci^+	affideranno	Italian
107			glie	lo		presento	

Italian shows one final variation (107), which might seem to falsify our arguments. As shown in §6.2.3, however, *gli*₃₋₃ is not *gli*_{DAT} but has ‘locative’ properties, leading many speakers to employ *ci* in this position. Thus, the two contradictory situations of Spanish/Catalan *le/li* and Italian *gli* emerge naturally from the already determined properties of clitics and structure.

7.4.4 Romanian

Romanian has a full DAT_[±E] paradigm allowing it to express event-internal coincidence directly, and ‘freeing’ it from the PCC. The difference between surface-identical DAT_[−E] and DAT_[+E] can be seen in their doubling behaviour where recipient/possessor DPs (DAT_[−E]) must be clitic-doubled unlike all the examples discussed below, and their use in passives (§3.4.1), which do not license DAT_[−E], only impersonal and personal locatives i.e. DAT_[+E].

Table 255													
108	1	2	3M	3F	R[±PL]	4	5	6M	6F				
ACC	mă [mə] m [m]	te [te]	îl [ɨl] l [l]	o [o]	se [se] s [s]	ne [ne]	vă [və] v [v]	îi [ɨi] i [i]	le [le]				
DAT	îmi [imʲ] mi [mi]/[mʲ]	îți [itsʲ] ți [tsi]/[tsʲ]	îi [ti] i[ij]/[j]		își [ɨʃʲ] și [ʃi]/[ʃʲ]			ni [ni]		vi [vi]	li [li]		
109	mă	te	ne	vă	o	l	le	i	se				
mi		✓											
ți													
i		✓											
ne→ni		✓										✓	
vă→vi													
le→li		✓											
și		✓							RND				

Clitics follow a rigid pattern (110), including some adverbials (all phonologically clitics, Dobrovie-Sorin 2013), taking prosodically determined forms (108). Ditransitive (111), and *present*-verbs (112) show similar behaviour with D/A-order pre- and post-verbally. There is no 3-3-rule (112).

Grammars proscribe many combinations (109), however, Ciucivara (2009)’s large-scale acceptability study shows that all combinations are at least marginally acceptable (i.e. interpretable) to some people in some circumstances. Like Italian, Standard Romanian gives preference pre-verbally to $\downarrow mi+te$, however, $\% \dot{\iota}i+m\ddot{a}$ and even $\%i-m\ddot{a}$ are widely acceptable. Most importantly, all SG+SG combinations are fully acceptable post-verbally (111), although not everyone accepts (113, $\% \dot{\varsigma}i+m\ddot{a}$). Pre-verbal singleton 1/2/3.PL take identical forms *ne/vă/le* in ACC (114) and DAT (115). Whilst 1/2.PL clusters $ni_{DAT}+v\ddot{a}_{ACC}/vi_{DAT}+ne_{ACC}$ are degraded for some (120-121) particularly pre-verbally (118), this cannot be due to number, since combinations with 3.ACC.PL are acceptable (116-117). Case syncretic 1/2.PL clusters are strongly ungrammatical pre-verbally (119), but acceptable post-verbally. Feature-based analyses cannot explain such variation. We propose that Romanian is *grammatically* unrestricted, but prosodically circumscribed.

Table 256

110	$Nu_{NEG} \dot{\iota}i_{DAT}=l_{ACC}=a_{AUX} mai_{ADV} \dot{f}i_{PFV} dat_{Vt}$	I would not have given it to you anymore
111	Dăndu $\dot{\iota}i/i$ mă de nevestă, tata...	Giving me to him/you in marriage, my father...
112	I l/le-am prezentat	I introduced him/them to her
113	$\%Lu\ddot{a}ndu \dot{\varsigma}i$ mă drept martor,...	Taking me as a witness for himself,...
114	Ne/vă/le vede	He sees us/you _{PL} /them
115	Ne/vă/le dă bomboane	He gives candy to us/you/them
116	Punându-ni-le în brațe,...	By putting them in our arms,...
117	Ni/vi le-a recomandat	S/he recommended them to us/you
118	$\{^{??}Ni v/^{??}Li v\}$ a recomandat	...you _{PL} to {us/them}
119	$\{*ne v/*v\ddot{a} ne\}$ au pus în brațe	He put {you in our/us in your} arms
120	$^?Prezent\ddot{a}ndu-\{ni-v\ddot{a}/vi-ne\},...$	When introducing... {you _{PL} to us/us to you _{PL} }...
121	$^?Prezent\ddot{a}ndu-li-ne/v\ddot{a},...$...us/you _{PL} to them...

Singleton clitics (122, other than *o*) attach phonologically to V-AUX,²²⁰ where (CL+AUX) is pronounced as a prosodic unit with clitics in reduced form. Clitics do not attach

220 This section use the abbreviations V-AUX (vowel-initial) vs. C-AUX (consonant-initial) auxiliary verb, and V-LEX (vowel-initial) vs. C-LEX (consonant-initial) lexical verb.

phonologically to C-AUX. (CL_{DAT/ACC}) and (C-AUX) are pronounced separately with full forms (123, *îi*). (CL_{DAT}+CL_{ACC}) form a prosodic unit with reduced clitics, even if there are potential phonological hosts preceding (128). (CL_{DAT}+CL_{ACC}+V-AUX) are pronounced as one prosodic unit (124), whilst (CL_{DAT}+CL_{ACC})+C-AUX are pronounced separately (125). Infinitive (126), subjunctive (127), and negative imperative (128) constructions follow the same pattern.

Table 257

Table 257

	I/S	NEG	DAT	ACC	AUX	INT	PFV	V	DAT	ACC	AUX
Infinitive	a	(NEG)	DAT	ACC	Ø	(INT)	(PFV)	V			
Subjunctive	să		DAT	ACC	Ø	(INT)	(PFV)	V			
Neg. Imperative	Ø	NEG	DAT	ACC	Ø	(INT)		V			
Indicative	Ø	(NEG)	DAT	ACC	AUX	(INT)	(PFV)	V			
			DAT	Ø _i	V-AUX	(INT)	(PFV)	V		o _i	
Conditional	Ø		DAT	ACC	AUX			V			
			DAT	Ø _i	V-AUX			V		o _i	
Conditional2	Ø						V	DAT	ACC	AUX	
Gerund	Ø	(NEG)				(INT)		V	DAT	ACC	
Imperative	Ø	NEG				(INT)		V	DAT	ACC	

122	(i=am) _o [dat un cadou	I/we have given him a gift	CL-Reduced	V-AUX ²²¹
123	îi (voi) _o [da un cadou	I will give him a gift	CL-Full	C-AUX
124	(ți=l=am) _o [dat	I/we have given you it _{MASC}	CL-Reduced	V-AUX
125	(ți=l) _o (voi) _o [da	I will give you it _{MASC}	CL-Reduced	C-AUX

126	a nu (ți=l) _o [trimite	Not sending it to you	Infinitive	
127	M=a rugat să nu (ți=l) _o [trimit	He asked me not to send you it	Subjunctive	
128	Nu (mi=l) _o [trimite	Don't send me it!	Negative imperative	

129	văzînd	Seeing	Gerund	Positive imperative
130	Văzîndu]-mă/-i	Seeing me/them		
131	Trimite]=(mi=l) _o	Send me it!		
132	Trimite]=l/*îl	Send it!		
133	dați [dats ⁱ]	Give!		
134	dați-l [dâtsil]	Give it!		

135	(l-aș) _o cânta~cânta-(l-aș) _o	I would sing it	Conditional	
-----	---	-----------------	-------------	--

136	O (voi) _o [trimite	I will send her		
137	O [aud	I hear her		
138	Eu <*o> am _{AUX} adus] <o>	I have brought it		
139	o [am _{LEX}	I have it _{FEM}		
140	Mîncă=(o=ar) _o mama	Mother would eat it _{FEM} /her		
141	cîntă+o→cînt[-o]	Sing it!		
142	șterge+o→șterg[əo]	Beat it!		
143	treceți+o→treceți[io]	Pass it!		

221 Data from Dobrovie-Sorin (2013).

Enclitics exhibit similar arrangements of clustering vs. extra-metrical singletons. Gerunds and imperatives are ‘filled out’ to maintain the prosodic boundary: plural imperative asyllabic *i* (133) becomes syllabic (134), whilst gerunds (129) receive syllabic *-u* before consonant-initial and semi-vowels clitics (130). In this prosodically strong position, insertion of *î-* is unnecessary (132). Verb and clusters are pronounced separately (131), whilst extra-metrical singletons are re-syllabified at higher levels of prosody, usually conjoining with the verb but also following words [arată=mⁱ] omul~arată [mj=omul] (Popescu 2000:158).

*O*_{ACC.FEM.SG} is exceptional, occurring before C-AUX (136) and V-LEX (137), but not V-AUX (138). In such cases, it follows the verb, from where it obligatorily modifies preceding vowels (141-143). This is prosodically, rather than phonologically, determined since *o* can appear before identical V-LEX HAVE (139). The past indicative (*am/ai/a...*), optative (*aş/ai/ar...*), and future (*voi/vei/va...*) auxiliaries may display inversion where clitics and auxiliary maintain their relative positions (135). Note that *o* can precede V-AUX post-verbally (140). These structures are archaic, but emphasize that (CL_{REDUCED}+CL_{REDUCED}), ((CL_{REDUCED})+CL_{REDUCED}+V-AUX) are units, in opposition to (CL_{FULL})+(C-AUX/LEX).

Boundaries exist between clitic-field and verb, filled out where necessary.²²² Within the clitic-field, the major determinant of acceptability is formation of appropriate prosodic units. Whilst hiatus exists in the lexicon under stress ('vi.e [ˈvi.e]~[ˈvi.je]), or morphological composition ([ˈre.a.na.li.zʌ] ‘re-analysis’), it is strictly avoided within inflections and CG (Chitoran 2002:§4.4). Variations in availability pre- and post-verbally reflects differences in strong/weak prosodic positions.

²²² Similar effects may be found when extra-metrical material is re-syllabified at higher levels of prosody (heavily influenced by speech rate) but existing feet/boundaries are always respected e.g. clitics already in groups cannot undergo optional phonetic cliticization (Popescu 2000:157-159).

Underlyingly, DAT.PL *ne/vă/le* possess mid vowels as shown when extra-metrical, and their glides before *-o/a*. In clusters, the vowel is required to raise before consonants and *i* which itself semi-vocalises (Gerlach 2002). Pre-verbal combinations of *ne/vă/le* are unacceptable because the first vowel ‘should’ raise producing *ni-vă/ni-le/vi-ne/vi-le/li-ne/li-vă* (which are acceptable to some²²³) just as *ne/vă/le+i→ni/vi/li+j*. Post-verbally, the same situation holds for some speakers. For others, strong-position inhibits raising, leading to *ne-vă/ne-le/vă-ne/vă-le/le-ne/le-vă*. Acceptability tests will, therefore, always return variable acceptability for such post-verbal cases: *ne-le* speakers voting down *ni-le* and vice versa, whilst both decry pre-verbal *ne-le*. Such variations do not reflect morphological legitimacy and number is only relevant in so far as syncretism is restricted to the plural.

Underlying *-i* operates differently. In a language-wide process of word-final high vowel desyllabification (Alkire & Rosen 2010:§10.1.8), */i/* forms glides following (144) or preceding (145) vowels and secondary palatalization gestures following word-final consonants (146), but is retained before consonant-initial inflections (147). Clitics in *-i* follow suit. In weak pre-verbal position, word-final *-i* of singletons obligatorily reduces causing insertion of initial-*î* [i̯] which becomes the syllable nucleus (i̯s^j-----).²²⁴ In clusters, they retain *-i* before consonants (*mi-te*) or form glides before vowels (*mj-o*). In strong post-verbal position, *î*-insertion is impossible because the verb-final weak position has been already filled if necessary. Plural syncretism blocked application of these rules for *ne/vă/le*.

223 Other factors such as general unavailability of *vă_{ACC}* (see below) also come into play.

224 [i̯]s are commonly treated as support vowels added to ensure prosodic minimality (e.g. Chereches 2014). Monachesi (2005), however, posits multiple allomorphs, whilst Popescu (2000) treats such clitics as containing underlying empty morae e.g. 3.SG.M.ACC */_μl/→[i̯l]*.

$O_{3.FEM.SG.ACC}$ must be expressed. From strong position, it modifies preceding vowels to form complex nuclei (150), retaining strong position and morphological content. V-AUX also takes strong position affecting preceding vowels, but whilst [ɔ] exists in many speakers speech, proclitic *o* never reduces, rather it appears after the verb. From post-verbal strong position, *o* obligatorily modifies preceding verb-final or clitic vowels (141-143). In this position, *o* is unaffected by V-AUX (140, now in weak position). In contrast, $\hat{n}_{3.SG.DAT}$ and $\hat{n}_{3.PL.MASC.ACC}$ convert to [j], fitting all positions/combinations. Similarly, *l* (124-128). Both take *î-* in isolation (123) for the same reasons as the (*i*)*mi*-type.

Table 258

144	$da_{GIVE}+i_{2.SG} \rightarrow da_j$ ‘you give’	Phonetic Rules
145	$ace\text{ș}t_{THESE}+a_{DEF.ART} \rightarrow ace\text{ș}tja$ ‘these _{MASC.PL.DEF} ’	
146	$lup_{WOLF}+i_{M.PL} \rightarrow lup^j$ ‘wolves’	
147	$lup+i_{M.PL}+lor_{GEN/DAT} \rightarrow lupilor$ ‘of/to the wolves’	
148	$karte_{BOOK}+a_{DEF.ART} \rightarrow kartea$, ‘the book’	
149	$fat\Lambda_{GIRL}+a_{DEF.ART} \rightarrow fata$ (*fat Λa), ‘the girl’	
150	$ne-O_{3.FEM.SG.ACC}/ne-am_{AUX}/ne$ cumpărăm	
151	arat-ă [mə]/*[m]!	See me!
152	[mə]/[m] arăt	I see myself
153	<*ti> ne recomandă <ție>	He recommends us to you
154	[tʃi.nə ar] face asta?	Wer würde das tun?
155	$tsi_{2.DAT}+atsi_{2.PL.AUX} \rightarrow tsj+ats^j$	
156	$te_{2.ACC}+atsi_{2.PL.AUX} \rightarrow t\text{e}+ats^j$, $\sqrt{mi}+t\text{e}+ats^j$ etc.	
157	{mi/ți/mi ți-i Ø _i } aduce împăratului _i merele _i	He brings the apples _i to the king _i
158	{mi/ți/mi ți-Ø-l} aruncă vrăjitoarea peste șapte codri!	The witch threw it over seven woods
159	<i>Te/se~ne/mă~vă</i>	<i>Grammar</i>
	<i>Te/ne/se~mă~vă</i>	<i>Standard Usage</i>
	<i>Te/ne/se/mă~vă</i>	<i>‘ți-mă’ dialects</i>
	<i>Te/ne~se/mă~vă</i>	<i>‘se=[sə]’ idiolects</i>

Final *-e* glides before vowels (148), hence (150), but Λ deletes (149) causing difficulties for clitics $mă_{ACC}/vă_{ACC}/vă_{DAT}$. Post-verbally, $vă/mă$ does not reduce since it is prosodic-word final, which is treated as a rule (151). Pre-verbally, the vowel is also expressed, although regularly reduces under higher-level re-syllabification before vowels in speech (152). The centrality of *ă*’s status is confirmed by the fact that for some speakers, $se=[sə]$ also reduces to *s*’ and causes

similar difficulties (Avram 2001). Thus, $\text{mi-te} \sim \text{ti-mă}$ indicates nothing about cluster nor dative, but about the weakness of *mă*, different in dialect/idiolects which accept *ti-mă/i-mă/și-mă*. Whilst grammars present *ti-ne* as ungrammatical (153), Popescu offers (154) as unquestionable; as confirmed by both surveys. Thus, there is a singular problem of *-ă* but dialect/idiolect variation in where it is found (159). It is not surprising that statistical surveys show ‘marginal’ (i.e. some speakers do and some don’t) acceptance of many clusters (particularly pre-verbally).

Many other factors must be taken into account. Grammars note (155-156), and restrictions on triplets involving ‘ethical’ datives *mi/ti/(mi+ti)* which may precede DAT (157) or ACC (158), but not clusters. $V\check{a} \rightarrow v$ is proscribed in clusters which may be related to potential confusion with va_{AUX} . Such observations further highlight that restrictions are not feature-based.

When pairs/triplets cannot combine (through language-wide phonological processes) into the required prosodic units, they are considered questionable. Post-verbal strong position ensures that all initial clitics are realizable, whilst the open position left for second clitics facilitates (although does not guarantee) realization of clitics such as *mă/vă*. Our analysis, therefore, expects few restrictions on enclitic clusters, but many on proclitics where weak position restricts vocalic combinations and hinders *mă/vă* realization. A prosodic analysis, therefore, fits the data where a morphological one fails. The reason why 1/2 *appear* relevant is because this is where *ă* occurs. The error in associating the issue with person, is seen in the fact that it does not extend to 2.SG *te*, but may extend to 3.SG *se*. Adding this to our understanding of plural combinations, explains why §7.3.1’s survey indicated that statistically neither person nor person/number is significant.

The current situation is a recent development. *O* previously could precede pre-verbal V-AUX, clitics did not invert with AUX, *î-* was not inserted (Popescu 2000:190), whilst post-posing reflexives was common up to XX^c (Tiktin 1891). Giurgea (2013) shows that replacement of *nă/lă* by *ne/le*, which only took place in Daco-Romanian dialects quite late (ancient texts retain *nă/lă*), resulted from a process of ‘velarisation’ of *e* (*e*→*ă* after labials unless followed by front vowels) creating a *me~mă*_{1.SG.ACC} alternation and triggering the emergence of *ne/nă*_{1.PL}, *le/lă*_{3.PL} pairs (and probably **ve/vă*_{2.PL}). This leads to the different behaviour of final -e vs. -ă clitics, and different acceptability in strong vs. weak positions. As prosody changed, positions became (un)available to each clitic and consequently cluster acceptability changed. This does not reflect upon number/person features nor associated exclusion rules. Restrictions are *not* results of banning specific combinations, but reflect suitability of individual clitics (each with their own properties) for their intended position. There is no **o+am*, but rather **[o]_{WEAK}*, no **î+mă*, but **[ă]_{STRONG}*, etc.

As a BE_AT language, Romanian possesses personal DAT_[+E] clitics, making most combinations ‘grammatical’ (i.e. interpretable) with ‘acceptability’ as a separate property. In addition to RND (incorporating context-based (un)acceptability of double plurals) and pragmatic considerations,²²⁵ the key property which degrades combinations is inability to fit their prosodic environment. Whilst we have not provided explicit explanations for every case, an analysis where ‘unacceptability’≈‘rhythmic awkwardness’ fits the empirical facts better than feature-based ones. Rather than the PCC being randomly breached, Romanian displays absence of PCC (because of presence of personal DAT_[+E]) overlaid and obscured by complex prosodic/phonological factors.

²²⁵ Farkas & Kazasis (1980) propose numerous pragmatic forces (related to discourse prominence) which disfavour combinations, including **îi mă arată*, ‘(S/he) shows me to you’.

7.4.5 No PCC-Violations

We conclude that clitics (and their source DPs) in apparent PCC-breaches are *not* DAT_[-E], whilst apparent order reversals represent different constructions: NOM_[+R]+OBL_[-R]~OBL_[-R]+ACC_[±R]. Combinations with SE represent diverse constructions with subtly different meanings, reflected in differences of relative clitic position. Surface-order variation relates to meaning, not extra-linguistic impositions e.g. person-hierarchies. The range found is determined by language resources and context (Table 259).

Further language-specific properties overlay this arrangement e.g. Romanian prosody, or Spanish/Italian dialect differences in OBL paradigms. The result has been presented as gradience between *strong*-PCC and *no*-PCC, but these variations do not reflect upon the PCC; these are not DOC constructions.

Table 259

Availability of					
Clitics:	DAT _[+E] are Ø except for Romanian, OBL _[+E] often shows incomplete paradigms (absent entirely in French), whilst only some languages have impersonal locatives e.g. Catalan <i>hi_{IMP}</i> .				
Construction:	SE _{NOM} is restricted to a few verbs/classes in some languages.				
Appropriateness to					
Context:	Use of clitics requires their referents to be already discourse-salient and syntactically local.				
Speech-Act:	Some constructions or (prescribed) clitic uses may be considered inappropriate in formal contexts for which these verbs are frequently used.				
Meaning:	In selecting constructions, speakers highlight different situational properties in order to express their view of the matter e.g. SE _{NOM} emphasizes subject involvement/satisfaction.				
LOC		OBL _[+E]	DAT _[+E]	1+2	2+1
		Romanian		✓	%
		Spanish		%	✓
Italian				✓	%
French				✗	✗
				Non-D/A combinations	

Key to the more complex approaches is the need to explain *me-first*-PCC and *te-first*-PCC. These concepts are not relevant to the PCC since no two D/A animates may combine in possessive relationships. DAT_[−E], as found with ditransitives, or monovalent verbs in possessor-raising, remains subject to the PCC *absolutely* in all languages, whilst apparent exceptions are separate constructions.

Under DAT+ACC, the only ‘processes’ are 3-3-rules producing clitics: identical to 3.DAT (Italian *gli*), generally used for other purposes (Catalan *hi*, Spanish *se*), with unique forms (Old Spanish *ge*), or Ø. In Surmiran (Anderson 2005:243), whilst clitics may combine in non 3-3-contexts (160, note A/D-ordering), 3-3-clusters are ‘banned’ (161-163); ACC and/or DAT must appear as complements. We see this as the result of a 3-3-rule: 3.DAT+3.ACC→Ø+3.ACC. Combined with the possesional requirement *[±ANIM] [+ANIM], this produces what appears to be *super-strong*-PCC (§7.2), but in fact is no more than the combination of existing properties. There is no need for such specialised descriptions, nor gradients between them.

Table 260

160	Ursus <las> <ans> ò purto <las bulias> <a nous>	Ursus brought the mushrooms to us
161	Tgi dat igl matg a Gelgia?	Who is giving the bouquet to Gelgia?
162	?*Tgi igl la dat	Who it her gives?
163	?*Tgi l' igl dat	

Beyond D/A pairs, we have only been able to identify one ‘real’ restriction in all the languages surveyed. Ibero-Spanish **me+te* is quite robust and requires explanation, but it is not **me_{DAT}+te_{ACC}* and hence does not reflect upon the PCC, nor can be expressed in or explained by general feature-based processes (§7.3).

7.4.6 Old Spanish

(164) summarises clitic combinations found in CORDE with representative examples (166-176, Alba de la Fuente 2012). *Vos/nos/os* were commonly used as singular polite forms (Penny 2002:138). The shift from Old Spanish (XI-XV^c) to Modern Spanish (XVI-XX^c) saw competition between *tú* and *vos* as non-deferential singular. With *tú*'s supremacy, *vosotros* replaced *vos* in plural contexts with *os* as its clitic, restricted to plural referents. New deferential forms *usted(es)*<*vuestro/a(s) merced(es)* were established taking 3-person clitics *le(s)*_{DAT/ACC}.

Table 261

Forms						Functions				
164	me	te	nos	os	vos	165	1.SG	2.SG	1.PL	2.PL
me		0		26	68	1.SG		✓		✓
te	√ ²²⁶		72			2.SG	✓		✓	
nos		0		7	5	1.PL		✓		✗
os	96		0			2.PL	✓		✗	
vos	1		1							
Identity:										
None										
Partial										
Full										

		N	O	D	A	=Modern
166	Probadme que nos os burláis y yo os obedeceré (1627) Disprove to me that you are mocking us and I will obey you _{SG}		nos	os		nos+te
167	Heme aquí, do vos me arrimo (1550) Here I am, where I get close to you _{SG}		vos	me		te+me
168	Esperad, que no me vos podréis escapar (1512) Wait, because you will not be able to escape from me		me	vos		me+os
169	Llanto tengo en que me os bañéis, cabellos, para limpiaros (1652) I have tears in which you may bathe yourself (on me), hair to wipe yourself		me	os		me+te
170	Señor tio... nos vos mucho encomendamos (1454) Uncle.. I/we commend myself/ourselves to you _{SG}	nos	vos			nos/me+te
171	Días cansados, duras horas tristes,...en años de pesar os me volvistes (1535-1575) Tired days and hard, sad hours, you turned into sad years to me	os	me			os+me
172	¡Oh, benditas pajaricas,...no os me vais (c.1529) Oh, holy little birds,...do not get away from me!	os	me			os+me
173	Es bueno, replicó Micas, que os me llevais mis Dioses... (1703) It is good -replied Micas- that you take away my gods from/on me...	os	me			os+me
174	Pues para eso os me ha dado mi padre (1535-1622) Since it is for that reason that my father gave me to you _{SG}			os	me	te+me
175	Amiga, ¿es éste el cavallero que me os embió? (1512) My friend, is this the knight that sent you _{SG} to me?			me	os	me+te
176	Amiga buena, bendita sea la ora que vos Dios... vos nos dió (1300-1325) My good friend, blessed be the hour in which God...gave you _{SG} to us			vos	nos	te+nos

226 *Te+me* is found in profusion throughout the corpus.

(166-169) are OBL+SE_{MID}, where indirect participants are not recipient/possessors but affectees of subject-internal ongoing processes. (170-171) are personal (170) and inanimate (171) anticausatives (SE_{ANT}+OBL). (172) represents 7 cases of SE_{ANT}+motion-verb with OBL=source. Such personal locative use is still found in Spanish dialects (§3.3) but lost in Ibero-Spanish. (173) is SE_{NOM}+OBL, whilst (174-176) are residual pre-PCC transitive uses.

Presenting the data by function (165, see modern equivalents) rather than form (164), reveals no person/number restrictions except double plurals. Excluding D/A clusters lost as the PCC developed, all cases include reflexives, matching the Modern Spanish pattern, where X_[+R]+Y and X+Y_[+R] ease interpretation, whilst X_[-R]+Y_[-R] may lead to ambiguity and is avoided.

Whilst *te+me* is found in profusion, technically possible *me+te* is not; all cases use *os/vos* for singular referents. Bello & Cuervo (1960) notes that *os+me* was common until XVII^c. (169) represents the last of 26 relevant cases of *me+os*. Thus, *os+me/me+os* stood for modern *te+me/me+te* whilst *os* served as deferential 2.SG/PL, disappearing when it specialized as non-deferential 2.PL. Similarly, cases of singular *vos* (176). In Modern Spanish, *os+me/me+os* are highly restricted; the formal contexts of *recomendar/presentar* favouring *le(s)=usted(es)* over *te/os*.

The previous existence of these combinations, indicates that lack of *me+te* surface-forms was not due to person/number interaction (*1.SG+2.SG), but correlated with speaker ability to show deference to their interlocutor. *Te*_{OBL/ACC} is considered insufficiently deferential. Indeed *te*'s most frequent use is with positive imperatives which actively shows lack of deference; polite usage requires *le(s)*. Deference is only necessary, however, for non-subjects, making *me*_{NOM/OBL}+*te*_{ACC} questionable, but *te*_{ANT}+*me*_{OBL}/*te*_{NOM}+*me*_{ACC} acceptable.

This may, therefore, reflect *convention* which, unlike grammarian **me+se*, is agreed within each speech community. In English, distaste for hiatus and subsequent insertion of palatal glides in ‘I [y]and You’, led to its proscription in favour of ‘You and I’. The rule did not cover accusative ‘me and you/you and me’ since neither created the same dissonance. This became a matter of ‘politeness’ rather than euphony, such that ‘me and you’ also became proscribed. Similar restrictions are found across Romance. In Occitan, disjoint subject/complement pronouns²²⁷ are always ordered 1»2»3 when conjoined (177, Romíeu & Bianchi 2005:203). Italian *io* does not have to follow other coordinated (pro)nouns (178), however, “*tu ed io* seems to be the preferred order in formal language” (Maiden & Robustelli 2000:115). Moreover, whilst 1.SG *io* may be conjoined with (pro)nouns (179), it is more commonly expressed with *con* (180). Similarly, Argentinian Spanish (Butt & Benjamin 1994:127).

Table 262

177	Jo e tu, a jo e a era	I and you, to me and him
178	L’abbiamo fatto io ed te	We did it, you and I
179	Io e Giulio studiavamo il francese insieme	G. and I were studying French together
180	Studiavamo con Giulio il francese	Lit. We studied French with Giulio

Latin-American dialects developed different deference rules, and *me+te* does occur. Cuervo (2003) documents many examples including constructions e.g. 4-clitic clusters (§7.5.1), rarely found in Ibero-Spanish. On this basis, the restriction is *not* upon **me+te*, but non-deferential *te* and constructions capable of supporting the pair. This places the restriction beyond local syntax/morphology. It operates at the same level as personal-*a* which also arose as choice of deferential clitic declined, making any combination of unbound pronouns (e.g. *presentó a_{ACC} me a_{LOC} te*) acceptable in contrast to clitic combinations where deference cannot be shown.

227 i.e. where, being in the same case, unbound pronouns have the freedom to change order.

7.4.7 PCC Conclusions

Although, the effect becomes strikingly apparent with clitics, this restriction type is not a clitic-specific property and does not operate at the clitic~clitic level. As in the discussion of appropriateness vs. person-sequencing (§7.3), clitics merely reflect wider semantic/syntactic selections. Whilst we may not have given a simple and absolute explanation for **me_{NOM}+te_{OBL}*, we can (like 3-3-rules) justify separating it out as a distinct property over-laid upon an otherwise simple system. Introducing greater complexity into that system (e.g. unsupported feature manipulation) simply leads to greater obscurity and error.

7.5 Exclusions

This section reviews the full range of clusters available in Latin-American Spanish which shows the widest range of combinations and lacks complications introduced by swapping and **me+te*. It is evident that the reality of ‘exclusions’ bares little resemblance to the complex proposals of previous models.²²⁸

7.5.1 4-Clitic Clusters

4-clitic clusters exemplify the effect of RND most clearly. Taking an agentive verb (*SE_{NOM}*) and an acceptable DAT/ACC pair (thus ending in *lo*), Table 263 permutes the clitics under *OBL_[¬R]* and *DAT_[±R]*. This leaves 10 potential cases: ‘✕’ marks breaches of RND. The remaining possibilities are the only combinations acceptable to speakers who use these complex patterns (Cuervo 2003). Note that *te+se+lo* is legal in (181), but not (188), even though functions and positions are identical. Grammaticality depends upon their relationship to SUBJ. Form-oriented systems cannot make such distinctions, banning both due to *te+se*. Case-models, however, achieve total accuracy with no clitic-specific mechanisms. The patterns exemplified in (182) and (184) appear to be very marginal, but see (205 and §7.5.3).

²²⁸ Tables are restricted to combinations of singular clitics in order to save space.

Table 263

$N_{[+R]}$	$O_{[-R]}$	D	A		N_i	O_k	D_j	A_l	
me					me _i	te _k	se _j	lo _l	llevé _i
te	x	x	lo		te _i	le _k	te _j	lo _l	llevaste _i
se					se _i	me _k	se _j	lo _l	llevó _i
					se _i	te _k	me _j	lo _l	llevó _i
					se _i	me	se X	lo X	llevó _i
					se _i	te	se X	lo X	llevó _i
					se _i	le X	me	lo X	llevó _i
					se _i	le X	te	lo X	llevó _i

$\left. \begin{array}{l} 181 \\ 182 \\ 183 \\ 184 \\ 185 \\ 186 \\ 187 \\ 188 \\ 189 \\ 190 \end{array} \right\} P_i \text{ took}_i \text{ it}_l \text{ away from } P_j \text{ on } P_k$

Harris (1994) questioned why there is a maximum of four clitics but most clusters are smaller. The maximum reflects the four positions, whilst the properties of NOM/OBL, RND and PCC result in a natural frequency distribution favouring smaller clusters. Since breaches of RND increase in likelihood with the number of clitics, the number of legal clitic patterns decreases with size of cluster. Person-models can offer no insight in this area.

7.5.2 Function, *not* Form

(191-199) show some of the interpretations possible for *me+te*. Either element may be ungrammatical depending upon its function, reflected in its position and subject co-reference; i.e. $NOM_{[+R]}$ and $OBL_{[-R]}$. The grammaticality and meaning of each identical pair varies based on each clitic's *function*. By treating all *me* the same, controlled by the same exclusion rules, grammatical cannot be separated from ungrammatical.

Table 264

	N	O	D	A	
191	me _i	te _j	se _{spur}	lo _i	I _i took _i it _j away from him _k on you _j
192		me _k	te _i	lo _i	You _i took _i it _j for yourself _i on me _k
193		me _i	te _k	lo _i	He _i took _i it _j from you _k on me _i
194	me _i		te _k	lo _i	I _i took _i it _j away from him _k
195		*me _k	te _j	lo _i	llevé _i
196	*me _i	*te _i	se _{spur}	lo _i	llevaste _i
197	*me _i	te _i	se _{spur}	lo _i	llevó _i
198	*me _i		te _k	lo _i	llevaste _i
199	*me _i		te _k	lo _i	llevó _i

$\left. \begin{array}{l} 195 \\ 196 \\ 197 \\ 198 \\ 199 \end{array} \right\} \begin{array}{l} *te \\ OBL=SUBJ \\ *me \\ NOM \neq SUBJ \end{array}$

What prevents any pattern surfacing is *not* inter-clitic mechanisms of exclusion/order, but rather, whether *each* clitic is *interpretable* in its position relative to the verb. Acceptable sequences are simply multiple clitics, each of which can be simultaneously interpreted in an acceptable way. Such an analysis is impossible in a person-model where a clitic’s validity is determined in reference to its neighbour, regardless of the function of either.

7.5.3 Delimiting the PCC

Permuting variations (200) for NOM_[+R] with OBL_[−R] and SUBJ (for intransitive verbs) shows that PCC restrictions do not apply within the upper clitic-field; *le+me* etc. are unavailable simply because *le_{NOM}* does not exist. Permuting variations (201) for OBL_[−R] with DAT_[±R] and SUBJ (for transitive verbs) shows that PCC restrictions do not apply across upper~lower field boundaries; *le+me* etc. are acceptable, if rare. Thus the PCC is only responsible for banning *le_{DAT}+me_{ACC}* etc. within the lower field. Person-models cannot delimit the action of a constraint leading to incorrect results; only case-models can provide a structural explanation of this behaviour.

Table 265

200

[IP]	N _[+R]	O _[−R]]
	i	j	
	me	me	
	te	te	
	se	le	

N _i	O _j	D	A		
me _i	te			morí _i	P _i up and died _i on P _j
	le				
te _i	me			moriste _i	
	le				
se _i	me			murió _i	
	te				
	le				

RND removes pairs of exact (e.g. *me+me*) and intersecting (e.g. *nos+me*) identity (§2.3.1). However, it allows 3-3-combinations where two referents can be distinguished. Such distinctions appear to be syntactic rather than referential. In (203), *le+le* is unacceptable even when referents are distinct e.g. ‘don’t shout at him on her’. As with 3-3-rules, only reflexive

vs. non-reflexive is sufficient distinction (204). Cases of *le+lo* may also be semantically distinct (*lo*=animate, *le*=animate). The combination is acceptable when 3-persons are separated (205), but otherwise is only marginally so in a clearly contrastive context (206). By default, it is interpreted as (207) where the spurious-*se* rule would be invoked. Whilst this oddity underlines the need for a more formal definition of ‘identity’ in RND, it offers no evidence for clitic-specific exclusions, since MCs cannot access the information required to make these even more delicate choices.

Table 266

201

[IP]	O _[±R]	D _[±R]	
	j	k	

me	me
te	te
le	se
	le

N	O _j	D _k	A _i	P _i gives _i it _i to P _j on P _k
	te	me _i	lo _i	doy _i
		le→se		
	le	me _i	lo _i	das _i
		te		
	me	te _i	lo _i	das _i
		le→se		
	le	me	lo _i	da _i
		te _i		
	me	te	lo _i	da _i
		le→se _i		
		se _i	lo _i	da _i
		me		
	te	le→se _i	lo _i	da _i
		se _i		
	le	me	lo _i	da _i
		te		
		se _i ✗	lo _i	da _i

← Only two 3-persons allowed!

		N	O	D	A		
202	No		me _k	le _j		grite _i	Don't shout _i at him _j on me _k /*him _k
203			*le _k				
204			le _k	se _j		abrió _i la puerta _i (a María) _k	The door _i opened _i on her _k (on M.) _k
205		me _i	le _i	te _k	lo _j	llevé _i	I _i took _i it _j away from you _k on him _i
206		te _i	le _k	→	lo _j	llevaste _i	?/*You _i took _i it _j away on him _k
207		te _i		se _k	lo _j	llevaste _i (lelo→selo)	You _i took _i it _j away from him _k

7.5.4 Proscriptions

Previous approaches define models suitable for one (highly idealised) range of usage and then attempted to ‘shuffle’ its constraints to match other usages. The current model is able to handle the full range of clitic clusters available in all dialect/idiolect variations (even if not everyone uses every one of them), to which further constraints for those who deselect certain cases may be added, *if necessary*. Thus, we start with a single open model which can be further constrained, rather than an indeterminate series of restricted models.

Since case-models already deal with ungrammaticality, we are concerned here with licit forms which are simply avoided by individuals, because they *feel* them to be less *usable* than others. It is doubtful whether anyone speaks standard dialects as defined by official grammars, given that use of OBL in everyday unguarded speech is so frequent as to require explicit proscription (§1.2.3). In reality, speakers are well aware of the potential for ambiguity and employ it in jokes:

-Mamá, mamá, me se cae la baba.
-No hija, será “SE ME”.
-No mamá, te juro que es baba.

M., M., the baby is drooling on me.
No daughter, that’s “se me”.
No M., I swear to you that it is the baby.²²⁹

Such cases illustrate that people can recognise and successfully parse these forms, even if they ‘disapprove’ of them. One might think in terms of speakers switching between multiple register-based grammars each with different (or differently ordered) rules, but it is much simpler to talk about a production restraint over-riding the same model, operating at (semi-)conscious levels e.g. ‘transitive constructions using OBL are avoided in well-educated

²²⁹ <http://www.blogdechistes.com/chiste/me-se-cae-la-baba.htm>. (‘BLOG DE CHISTES » Me se cae la baba | Los mejores Chistes cortos’ 2012).

society.’ As Russi (2008) notes, it is common for Italian speakers to believe that they follow grammarian rules but fail to do so in practice i.e. when semi-conscious *control* is inactive, during unguarded speech.

As illustrated throughout this work, ‘grammatical’ restrictions are remarkably few. Along with appropriate context, the central factor which determines usage is whether the sequence is communicationally effective, the evaluation of which operates on two levels: (1) does the form perform the necessary social function; will its use make the listener think less of the speaker, or show solidarity with them? (2) Can the speaker be sure that the listener will follow his/her intent. If not, a different construction will be used. The result of (1)+(2) is that people who rarely experience these forms tend not to use them (even if they know that they are possible), because they imagine that their interlocutor will feel similar issues in decoding/accepting messages so presented. The ‘missing’ constraints are, therefore, not grammatical processes but elements of communicative competence based on the speaker’s encyclopaedic knowledge of his language and audience. They are (semi-)conscious choices rather than grammatical impositions. Models attempting to manage such complex choices by morphological movement/exclusion cannot cope with the range of subtle choices made in everyday speech.

Whilst formula such as $*me_{\text{NOM}}+te_{\text{OBL}}$ are useful shorthand descriptions, they should not be seen as defining processes. One result of a case-model is that $*X+Y$ style negative exclusions, don’t have any place in an adequate clitic model.

7.5.5 Re-Evaluating RND/PCC

Whilst $[\pm E]$ has been presented as a simple contrast describing the relationship between participants or groups thereof, it has its basis in set theory. A dominant partner in a relationship is construed as container of its subordinate partner $[-E]$, or a place of reference for that partner i.e. significant coincidence $[+E]$. This represents two of the four relationships available to sets (208-215).

Table 267

	Singular		Plural		Relationship	Property
Disjoint	208	{a} {b}	209	{a a a} {b b b}	$A \neq B$	$[+E]$
Subset	210	{a {b}}	211	{a a a {b b b}}	$A \neq B$	$[-E]$
Union	212	{a b}	213	{a a a b b b}	$A \neq B, A+B=AB$	New Item
Intersection	214	{a [?] b}	215	{a a a [?] b b b}	????	Impossible

In (208/209), *b* is identified as the object in *a*'s vicinity. In (210/211), *b* is identifiable as a distinct item but part of, and identified by, *a*. Logically, disjoint/subset *a* and *b* must be unique. It is impossible to be disjoint from oneself, or part of oneself but independent. If separated from *a*, *b* part becomes a disjoint item (see mortician examples, §3.2).

Union (212/213) creates a new single set e.g. I+you→we. For intersection, the question arises of what goes in the overlap, such that it is part of *a* and *b*? (214) might be possible for conjoined twins, where the intersection indicates the areas of their bodies shared. Otherwise, it is meaningless. Similarly, (215) cannot exist. It is impossible in these cases to identify what is being described, and hence impossible to find their referents.

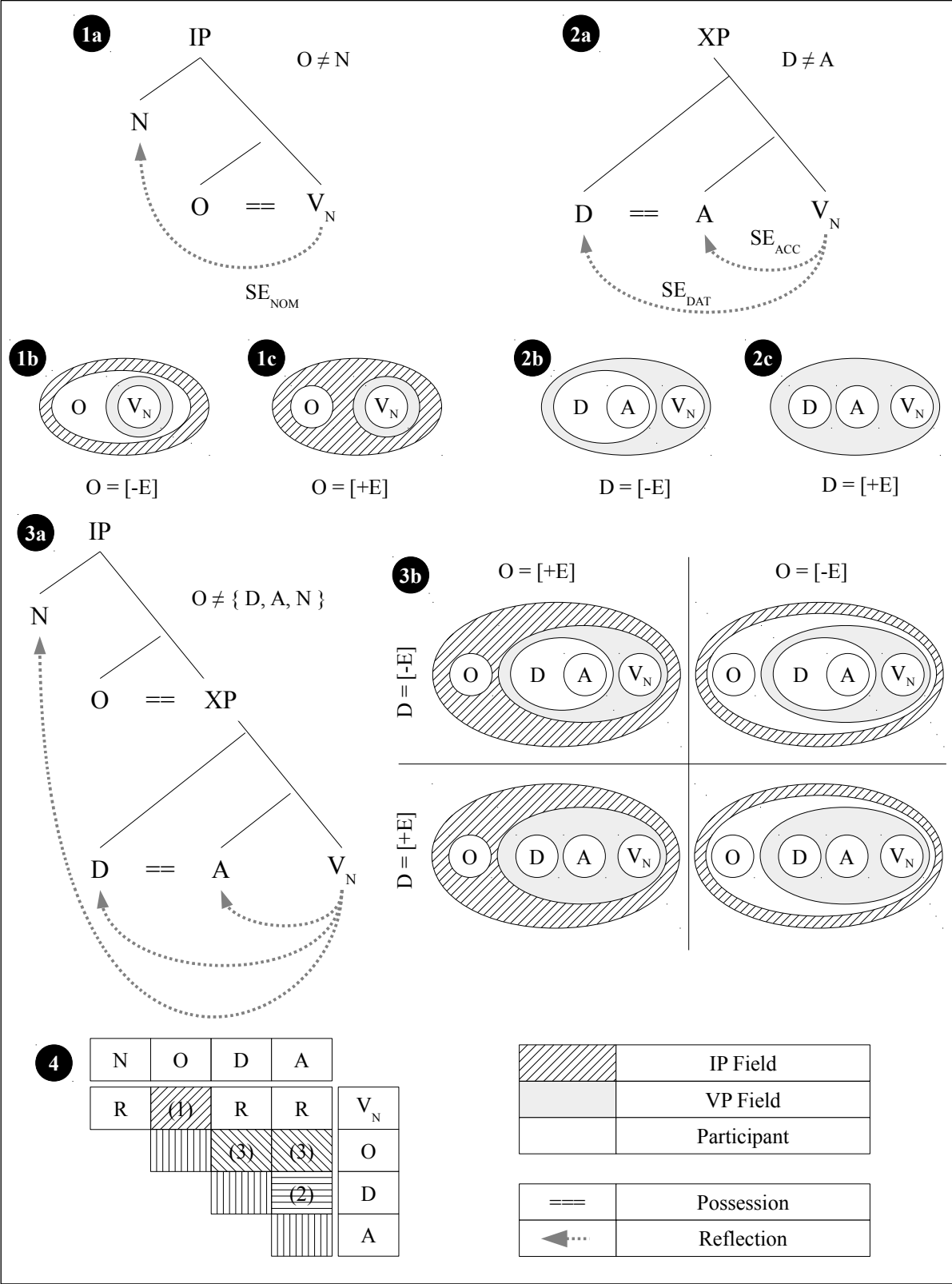
Thus, we are limited to two relationships ($[\pm E]$) which guarantee non-equivalence between two entities. Below we show that this guarantees uniqueness across our syntactic structure,

harking back to Strozer's exact~intersecting identity (§2.3.1), and linking it to the PCC as we have come to understand it in this chapter.

The final diagram presents the structural relationship between clitic-fields and their components. In IP (1a), a possessive relationship holds between OBL and V_N (which references the subject as participant). The two possible relationships (possession~coincidence) mirror the OBL division between sympathetic~setting datives (§3.3). As the set diagrams illustrate, whether in part-whole relationship (1b) or coincidence (1c), OBL cannot be the same as NOM. It follows that V_N will never 'reflect' onto OBL, but may onto NOM, creating SE_{NOM} . For VP arguments (2a), the relationship holds between DAT/ACC, guaranteeing participant independence, but since V_N is independent, it may 'reflect', creating $SE_{DAT/SE_{ACC}}$.

In full transitives (3a), OBL bares the same relationship with XP, and therefore is distinct from its participants; $OBL \neq DAT$, $OBL \neq ACC$. Since $DAT \neq ACC$, $SE_{DAT} \neq SE_{ACC}$. Combining these set relationships in (4), shows that all referents must be unique. RND is an emergent property of the two possessive relationships, and their relationship to each other, including restricting reflection to $SE_{NOM}/SE_{DAT}/SE_{ACC}$; thus, lack of SE_{OBL} is also emergent.²³⁰ Gallo-Romance SCL's are subject coreferent but not reflexive and can combine with SE_{DAT}/SE_{ACC} (§2.1.3). The only restrictions not directly encoded in structure are $SE_{NOM} \neq SE_{DAT}$ and $SE_{NOM} \neq SE_{ACC}$, i.e. V_N may only 'reflect' on a unique participant. Since explanation of this would lead to theory-specific considerations, we leave it as a stipulation which we believe to be acceptable to any theoretical approach: V_N 's features can only be interpreted once.

230 Overlapping identity e.g. *me+nos* would require intersections across participants and/or structures.



Whilst *DAT_[±ANIM]+ACC_[+ANIM] describes PCC effects, consideration of animacy obscures the central relationship: possession~coincidence. Whilst ACC_[±ANIM] may occur as *effected* objects, appearance under DAT_[−E] is determined by a participant's ability to possess ACC. DAT_[+ANIM] represent archetypical possessors, whilst DAT_[−ANIM] is only possible in part-whole relationship with other [−ANIM] participants. ACC_[−ANIM] represent archetypical possesseees, but not ACC_[+ANIM] which cannot be in part-whole relationship with DAT_[±ANIM], merely coincident. Expression of any of these relationships *as complements or clitics* is determined by each language's lexicon. In most languages, DAT_[+E] clitics are Ø resulting in them being inexpressible as clitic-pairs, unless the language has impersonal locatives e.g. Italian *ci_{IMP}*, or personal DAT_[+E] clitics e.g. Romanian, where such relationships are expressible and the language (modulo prosodic effects) is 'free' from the PCC.

From this perspective, languages start with [+E,+SPEC], [+E,-SPEC], [-E,+SPEC], [-E,-SPEC] weak personal and adverbial pronouns and lose some in the process of grammaticalization into clitics, or later in the development of clitics within each language. Romanian, preserving dative case and remaining a BE-AT language retained [+SPEC,±E] i.e. personal possessive and coincident DAT/OBL, but lost [-SPEC,+E] i.e. so-called locatives. Most other languages lost [+SPEC,+E] i.e. personal coincident DAT leading to the PCC coming into being. Many further lost [-SPEC,+E] i.e. locative clitics as well. The development of the PCC, is therefore simply the loss of coincident datives during the process of grammaticalization from WPs to modern clitics (i.e. heads to functional projections).

Lack of a N/O possessive relationship determines that there will be no upper-field limitation upon pairs based on animacy (or rather possession), nor upon OBL's possession of the neuter

event. The D/A possessive relationship guarantees $*me_{DAT}+*te_{ACC}$, $*le_{DAT}+me_{ACC}$ etc. i.e. a ‘PCC’ restricted to the lower-field, unless the language has coincident-marking clitics. That the PCC operates across complements as well as clitics shows that it is structure which determines these ‘exclusions’ not clitic-specific rules/mechanisms. Indeed, structure (i.e. the possessive relationships encoded in it) removes the need for any ‘exclusions’ in this area. Animacy is not the PCC’s motivation, but a reflection of underlying properties of possession which determine, not only VP-centred PCC, but also clause-wide RND. Far from a complex additional property of language, the PCC turns out to be equally emergent from the possessive structures posited, whilst RND turns out to be a ‘description’ of the product of the two asymmetric possessive relationships. It operates across language, because it operates across the logic of construal.

Table 268

216 **Agreement**

Subject	Reflexive clitic	NID	IT	SP
[III, +SPEC, ±DEF]	[III, +SPEC, ±DEF, +R]	si	si	se
[III, −SPEC, ±DEF]	[III, −SPEC, ±DEF, +R]	si	ci	Ø

217 **3-3-Process: 3.DAT+3.ACC →**

Surmisan	Ø+ACC/DAT+Ø ²³¹	Ø	} Possible outcomes
Old Spanish	ge [3e]	Unique (later >se)	
Modern Spanish	se [se]	Identical to reflexive se	
Italian	gli	Identical to 3.DAT	

Given the findings of this chapter, we conclude that there are no clitic-related exclusion mechanisms in Romance. There are two methods to ‘create’ a ‘missing’ item: agreement and 3-3-rules. The classic example of an unexplainable ‘random’ exclusion is that of impersonal reflexives (216, §4.6.8-4.6.10). Under our approach, [+SPEC] subjects require [+SPEC] reflexives, and [−SPEC] subjects require [−SPEC] reflexives, each has a separate place in the clitic lexicon. The only complexity which we recognise is 3-3-rules, which follow from RND

231 e.g. Tgi { *igl+la/*la+igl/ /Ø+la } dat

Who gives it to her?

admitting two 3-person clitics with distinct referents (217). 3-3-rules may produce surface forms that are: identical to 3.DAT; unique; look like another form; or Ø (e.g. Surmiran, §7.4.5). Whether either are available and what form it takes is a matter of historical accident. Like the PCC, these are not exclusions, merely Ø entries at the intersection of the syntax related columns and the referent related rows.

7.6 Conclusions

This chapter has extended the semantic basis of RND (i.e. logical availability of only disjoint vs. subset relationships) to cover what have previously been considered separate mechanisms of number-incompatibility and the PCC. The initial difficulties presented by *present*-verbs and Romanian turn out to be, not exceptions, but strong evidence for the proposition, where the differences between languages follows from language-specific properties which may be determined independently of this particular phenomenon. Furthermore, several related phenomena such as ‘inverted readings’ and $\sqrt{se+le} \sim *le+se$ emerge from the model as the only possible result rather than difficulties which require explanation.

A simple table (as defined in Chapter 2) including Ø entries and weights is sufficient to define all orders and exclusions, without recourse to any clitic-specific mechanisms, except 3_{DAT}-3_{ACC}-rules. This makes the remaining **me+te* type highly specialized, but we can (like 3-3-rules) justify its isolation, showing that it is *not* amenable to explanation in terms of featural processes as previously presented. An analysis based on inappropriate use of a familiar pronoun in formal contexts seems, in our opinion, to have more potential and better fit the reality of the situation.

8 CONCLUSIONS

The introduction observed that the general trend has been to exorcise clitics from syntax, pushing their ‘inexplicability’ to some other module of language e.g. storing “unanalysable chunks” in the lexicon or creating complex intervening morphological buffers. This work has presented a model where syntactic structure reflects semantics mapping (almost) directly onto surface forms, putatively ‘stored’ groups of clitic+verb (e.g. ‘inherent’ reflexives or ‘idiomatic’ phrases) are compositional and transparent, the arrangement of clusters displays clear structure/order (rather than representing freely associated sets), and no feature-based operations and/or associated re-ordering are required.

8.1 Summary

In the model (diagram opposite):

- Everything that is common to Romance is in the top (semantics/syntax) and bottom (prosody) sections. Everything that is language-specific resides in the clitic lexicon. Languages possess different lexicons as a product of their historical development, whilst knowledge of that lexicon helps determine which constructions are chosen and presented to syntax. People do not use clitics which do not exist in their language, or patterns felt to be inappropriate to the speech situation. The latter effect skews frequency of particular combinations, rather than acts as an out-right ban.

- Contra concerns of lack of ‘economy’ in such a ‘repetitious’ lexicon, we treat each ‘box’ as featurally fully defined and explain duplication of form through syncretism. We argue that form and function must be separate in order to explain the fluid change across each of these dimensions which we see during Romance’s history (Chapters 2 and 6). Moreover, for every set or pair which looks the same and might be a suitable target for reduction, there exists at least one language which maintains surface distinctions between them. If we are to make meaningful cross-linguistic comparisons, we have to compare functions, not forms.

- There exists a **necessary** set of restrictions on combinations of participants within **semantics**, which are reflected through structure across the clause i.e. verb + complements/clitics, obviating the need for clitic-specific restrictions in morpho-syntax. The set analysis in §7.5.5, showed that semantics limits pairs of entities to two relationships ($\pm E$) which guarantee their non-equivalence. Recursively, this guarantees uniqueness across our asymmetric syntactic structure, providing the basis for Strozer’s intersecting~exact identity (§2.3), and explaining why two plurals may show contextual variability in acceptance. RND becomes a ‘description’ of the product of two asymmetric possessive relationships. It operates across language, because it operates across the logic of construal.

- The four case model removes numerous surface combinations from DAT/ACC to NOM/OBL, whilst recognition of swapping helps us see what is actually there, removing the need for complex processes. This allows for simpler explanations of truly DAT/ACC related phenomena, in particular the PCC, which we argue like RND

is not a clitic-related process, but reflects the DAT/ACC “possessive” relationship, including restrictions inherent in the subset[−E] vs. disjoint[+E] dichotomy. NOM/OBL are not in a relationship, whilst OBL’s relationship is with the event substructure itself, and thus shows no combinatorial limitations. Restriction of the PCC to DAT/ACC, therefore, emerges for the asymmetric possessive relationships. DAT/ACC pairs are limited by their ability to “possess” another: animates may possess items but not people, whilst places are unfettered. There is no gradient between non-, partial-, or full-PCC languages. The mechanism is the same across Romance (semantics/logic), only the (independently provable) availability of particular clitics in the language-specific lexicons changes (§7.4.5). Even the PCC’s development can be seen as a historical process of loss of personal-locatives from Latin in most of Romance, but preserved in Romanian (§7.4.4).

- In Chapter 7, we conclude that there is no exclusion mechanism, which we feel accords with Baker’s Paradox. People do not think impossibilities and try to express them, only to have an autonomous (schizophrenic) morpho-syntax ‘correct’ them. Indeed, such errors cannot be ‘corrected’; deletion merely increases the confusion. By definition, exclusions reflect non-experience and cannot be learnt.
- Swapping due to weight (also seen in complements) is a fundamental to understanding the development of Romance. As illustrated in the Provençal study (§6.5), we can follow weight development from one snapshot in time to another, watch its effects change as dialects diverge, and explain numerous synchronic conundrums set by previous investigations. The general trend due to phonetic erosion is towards loss of

weight, hence the Romance-wide trend from predominantly A+D to D+A, but it occurs in waves affecting sets of clitics at different times, rather than a simple change of a D/A parameter. Moreover, the process can reverse as in the Roergat dialect. The result of these processes is that modern dialects can now be divided on the basis of the relative weights of their clitics. For each combination of FUNCTION, each dialect creates subtle variations in SURFACE sequence which has previously been impossible to capture, and left to ‘free’ variation. In reality, their forms and sequences follow directly from historical syncretism and change of weight.

- RND predicts that the only ambiguity will be in 3-3-contexts, since this is the only case where two clitics may surface with the same 3-person clitics, as long as they have different referents. The process may be *described* as a simple mutation, where 3-3-outputs may be, identical to 3.DAT, unique, re-use another form, or surface as Ø i.e. whatever developed in the ‘OTHER’ position. In fact, the equation is not quite so simple: ACC may be effected and/or effect DAT e.g. when ACC cannot be focused (e.g. [–SPEC]). As (§6.11) showed, the triggering conditions may involve several variables; an area which requires detailed (contextualized) investigation.
- The result of a 3-3-context may be subject to swapping, such that weight effects must be removed to get back to the underlying structural sequence. In doing this, numerous complexities become surface obvious (e.g. Catalan, §6.4), removing the need for morphological buffers capable of featural processes, or even spell-out rules.
- Due to forms shared across paradigms and the presence of Ø’s different constructions may result in the same surface sequence of clitics. The same sequence often has (out-

of-context) more than one possible reading, But clitics (contra *autosyn*) can only be interpreted in context.

- These situations are often presented as ambiguous. In the current model, only 3-3-contexts can show real ambiguity and each language has a way to deal with this in its 3-3-rules. In all other cases, RND ensures that each item is unique. What remains is *vagueness* as illustrated for reflexive~reciprocal usages (§4.2.2), DAT/OBL (Chapter 3) etc. Vagueness is inherent in language. It must be modelled, and most importantly, it must emerge from that model rather be stipulated. The current model correctly predicts vagueness and where it may occur, matching real life usage.
- Speakers are facilitated by a close relationship between the symbolic sequence and the experience being communicated. Minimising the cost of “processing enrichment” is key to easy communication, which is why messages characteristically display motivational or “diagrammatic” iconicity: “we keep finding iconicity because there is no other way for a semiotic system to be created and used by human beings without a close fit between form [in our case sequence of forms] and function” (Slobin 2005:320).
- Semantic properties are iconically reflected in structure as a chain of affectedness (§2.1), guiding evaluation through its inner→outer sequence. Structure tells the listener that more than one option is available, whilst default strategies (over-ridden by explicit data) lead to selection of an appropriate schema. In limited cases, there will remain more than one possibility and limited (and correctable if necessary) differences in understanding will ensue (§3.5.2). If the speaker (simultaneously a listener) believes

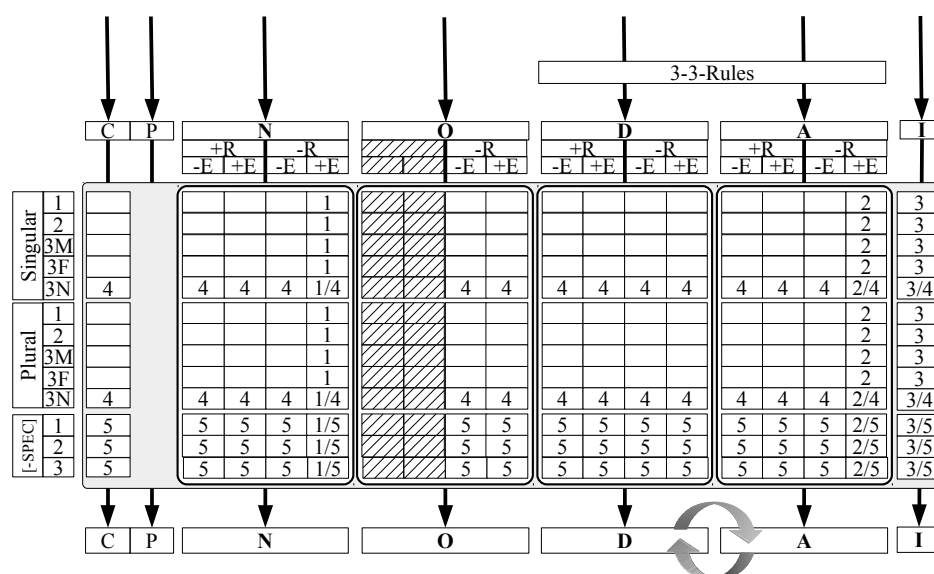
that mis-communication will occur, (s)he will select a different construction or add information which forces a particular reading. By definition, clitics refer to old/shared information. If that shared understanding no longer holds, e.g. the speaker needs to emphasize/contrast an element, it is reintroduced as a complement. The mere fact that a clitic is used indicates its low salience and semantic impact.

- Some clitics require more steps in their interpretation. It is easy to find the referent for most clitics e.g. *me* is always *me*. Some clitics (often referred to as ‘adverbial clitics’) require a further step in their interpretation. This is often confused with ‘idiomatic’ readings and the need for lexicalization. Like evaluation working through a fixed sequence of case, we argue that interpretation follows a fixed sequence of ever broadening semantic categories (§5.1.2). Whilst clitic properties remain constant, the most accessible topics change with discourse, hence interpretation follows context and identical phrases may give rise to several more or less idiomatic interpretations. This is impossible if meanings are lexically fixed.
- The combination of evaluation and interpretation sequences guides the Listener to interpret each variable as specific (e.g. a previously discussed place/a subset of a known entity) or ‘idiomatically’ (e.g. an abstraction such as the situation/a generic class of entity) in relation to subject or object. Under such a scheme, items *cannot* be freely ordered, and special/independent interpretation rules are unnecessary. Either approach would break the relationship which allows listeners to choose between specific~idiomatic readings, and evaluate who is doing what to whom. Contra García (§1.4.2), it is structure that allows transfer of meaning through such limited resources.

- The ability to deal with vagueness is a sign of communicational efficiency, limiting the need for repetition and explicit transfer of data. By virtue of such automatic inferences, increased explicitness signals variation from the norm. In ‘default’ contexts such explicitness becomes misleading to the listener. The gap between the correct default interpretation (denied by over-specification) and an alternative (demanded by inappropriate levels of specificity) causes a psychological dissonance often referred to as ungrammaticality. Most unacceptable usages are reasonable given an appropriate context, and therefore, should not be subject to ‘rules’ to ban them. In these cases, ungrammatical simply means inappropriate to context. Their inappropriateness is precisely because the listener expects to interpret the spoken message from context and minimal signals (§3.5.1).

8.2 Areas Not Covered

There remain ‘grey’ areas:



For the vertical zones (1-2), we have shown the need for this many contrastive categories, but when each is used (and its significance) is not clear. Similarly, the use of a ‘nominative’ class

of clitics (3) as found in Italian *si*_{IMP} and inverted questions in many NIDS (§2.1.2) is left for future research. For the horizontal zones (4-5), more detail is required in order to sub-categorise uses. This applies particularly to 3.Neuter which coalesces a range of properties such as \pm DEF, the mass~count distinction and ‘referentiality’.

This study has enabled us to identify regions of interest and the variables which must be considered. For example, at the current level of detail, we can justify isolation of 3-3-rules, and identify the variables which appear significant (§6.11), which (contra previous proposals) do not include *person*. To move forward, we need *more* detail; a survey which tests against the full range of variables in unambiguous contexts which help informants identify the *intended* communication and hence make their acceptability judgements meaningful. Otherwise, tests will continue to measure large and amorphous categories, rather than deliver clear insights. This needs to be carried out across Romance. Only with adequate volume of comparative data can we hope to spot the patterns underlying the phenomenon, rather than observe localised ‘descriptions’.

Whilst previous studies have offered numerous insights, their results (being expressed in differing models) remain disjoint. The most important feature of this work is that these explanations are offered through a single model, with a single representation of the clitic-lexicon shared across Romance. This opens that possibility of creating an online database allowing linguists to efficiently share information developed from corpora and specific studies. The model allows linguists to *rapidly* ‘fill in’ a table from *simple* activities, *predict* what will happen in complex cases, and *test* those predictions. By drawing together *comparable* evidence across potentially hundreds of clitic lexicons in a simple way, we can

focus upon areas where the model lacks detail. With synchronic and diachronic data, we have two orthogonal dimensions of contrast to constrain and validate our argumentation.

8.3 Conclusions

The approach taken in this work has been to build upon basic principles which we believe to be already present in the language:

- Independence of clitic form and function, as evidenced across time (‘overlap’ is the basis for reanalysis)
- Relations of objects along multiple dimensions (allows ‘spreading’ of forms by analogy and (over)generalization by learners).
- D/A swapping as found in complements (mirrored in the Romance-wide historical trend of clitic $A+D > D+A$).
- A coherent initial semantic graph, here represented in sets (we shouldn’t need exclusion mechanisms for logical impossibilities, which can never be experienced and learnt).
- Fixed evaluation sequence and interpretation consistent with that of complements.

Higher-order properties emerge from this base without stipulation or additional complexities. The model can be learnt through positive experience only, with uncorrected over-generalization and reanalysis leading to analogical processes i.e. historical change.

The model displays a direct relationship between semantic roles, syntactic case, and surface position, which holds across Romance, allowing language-specific detail to be fully

accounted for with a simple clitic lexicon. It defines a simple (although highly specific) structure for the clitic lexicon capable of displaying the processes of historical change found throughout Romance by simple well understood processes of phonological erosion, reanalysis and analogy. It defines a clear process of evaluation (in line with syntax) and interpretation (in line with semantics), resulting in no need for complex interpretation rules/mechanisms.

All putative exclusions emerge from the structure. There is no need to stipulate them as separate mechanisms. Beyond the swapping of ACC/DAT in the lower clitic-field, there is no evidence (or need) for inter-clitic movement or jockeying for position. Clitics surface in their syntactic position, which is an iconic representation of the underlying semantics. There is no template into which they are required to fit, or which has the ability to select, shuffle or delete them, nor indeed any movement which might require special syntactic rules or mechanisms. With the exception of 3-3-contexts, there is no evidence for featural processes, and even this may turn out to be simply a case of selection from an as-yet under-differentiated set of options.

This model does not delete grammatical forms, nor allow ungrammatical forms; although it does allow forms which might be unacceptable to some individuals/registers. It does not suffer from theoretical and practical problems such as transitivity or competition (since these are artefacts of imposing templates and/or person-ordering), whilst it provides natural answers to questions such as maximum and modal numbers of clitics in combinations. It provides a means to distinguish, and thereby analyse, differences such as agent vs. patient reflexives without stipulation or itemising them in the lexicon, whilst providing accurate coverage of the whole range of clitic combinations, without specialised mechanisms or stipulations.

Our analyses are less ‘explanations’ than ‘observations’ of properties which emerge unaided from the underlying model. Most importantly, these are properties that can be observed by learners and by such experience learnt. Under Occam’s Razor, the theory with the greatest coverage and least complexity should always be preferred. We opine that this model fits that description.

9 CORPORA

Catalan

Corpus del català contemporani
<http://www.ub.edu/ccub/>

French

BFM : La Base de Français Médiéval
<http://bfm.ens-lyon.fr/>

Corpus de Référence du Français parlé
<http://sites.univ-provence.fr/delic/corpus/index.html>

Corpus of spoken French
<http://www.llas.ac.uk/resources/mb/80>

Frantext
<http://zeus.inalf.fr/frantext.htm>

Italian

Asis Atlante Sintattico d'Italia
<http://asis-cnr.unipd.it/>

Banca dati dell'italiano parlato (BADIP)
<http://languageserver.uni-graz.at/badip/badip/home.php>

CORpus di Italiano Scritto (CORIS)
http://corpora.dslo.unibo.it/coris_eng.html

Corpus OVI : L'Opera del Vocabolario Italiano
<http://www.vocabolario.org/>

Libricino
<http://www.accademiadellacrusca.it/it/biblioteca>

ItTenTen10 – Corpus
<https://www.sketchengine.co.uk/ittenten-corpus/>

Portuguese

Tycho Brahe Parsed Corpus of Historical Portuguese

<http://www.tycho.iel.unicamp.br/~tycho/corpus/en/>

Romanian

Romanian corpus of newspaper articles

<http://www.cse.unt.edu/~rada/downloads.html#romainan>

Spanish

Corpus Del Español: 100 Million Words, 1200s-1900s

<http://www.corpusdelespanol.org/x.asp>

Corpus Oral de Referencia de la Lengua Española Contemporánea CORLEC

<http://www.lllf.uam.es/ESP/Corlec.html>

Corpus Oral y Sonoro del Español Rural (COSER)

<http://www.lllf.uam.es:8888/coser/>

Real Academia Española - Corpus Diacrónico del Español (CORDE)

<http://corpus.rae.es/cordenet.html>

Real Academia Española - Corpus de Referencia del Español Actual (CREA)

<http://corpus.rae.es/creanet.html>

10 BIBLIOGRAPHY

- Abeillé, A., and Godard, D. (1997) 'The syntax of French negative adverbs', in Forget, D., Hirschbühler, P., Martineau, F., and Rivero, M. L. (eds.) *Negation and Polarity: Syntax and semantics. Selected papers from the colloquium Negation: Syntax and Semantics. Ottawa, 11–13 May 1995*. Amsterdam: John Benjamins Publishing Company, pp. 1-28. Available at: <http://www.jbe-platform.com/content/books/9789027275899> (Accessed: 27th December 2014).
- Abeillé, A., and Godard, D. (2000) 'French Word Order and Lexical Weight', in Borsley, R. (ed.) *The Nature and Function of Syntactic Categories*, Syntax and Semantics vol.32. Academic Press, pp. 325-358.
- Abeillé, A., Godard, D., and Sag, I. A. (1998) 'Two kinds of composition in French complex predicates', *Syntax and Semantics*, 30, pp. 1-41.
- Ackema, P., and Schoorlemmer, M. (2005) 'Middles', in Everaert, M., and van Riemsdijk, H. (eds.) *The Blackwell companion to syntax*, vol. III. Oxford: Basil Blackwell, pp. 131-203.
- Ackerman, F., and Moore, J. (1999) 'Syntagmatic and Paradigmatic Dimensions of Causee Encodings', *Linguistics and Philosophy*, 22, pp. 1-44.
- Agostini, P. M. (1984) *L'usu di a nostra lingua. Grammaire descriptive corse*. Bastia: Edizioni Scola Corsa.
- Aissen, J. (2003) 'Differential Object Marking: Iconicity vs. Economy', *Natural Language and Linguistic Theory*, 21, pp. 435-483.
- Aissen, J., and Bresnan, J. (2004) 'Remarks on description and explanation in grammar', *Studies in Language*, 28, pp. 580-583.
- Al-Zahre, N., and Boneh, N. (2010) 'Coreferential dative constructions in Syrian Arabic and Modern Hebrew', *Brill*, 2, pp. 248-282.
- Alarcos Llorach, E. (1970) *Estudios de gramática funcional del español*. Madrid: Gredos.
- Alarcos Llorach, E. (1994) *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Alba de la Fuente, A. (2012) *Clitic combinations in Spanish: Syntax, processing and acquisition*. Doctoral Thesis. University of Ottawa.
- Albizu, P. (1997) 'Generalized Person-Case Constraint: A case for a syntax-driven inflectional morphology', in Uribe-Etxebarria, M., and Mendikoetxea, A. (eds.) *Theoretical issues on the morphology-syntax interface*. Donostia: Gipuzkoako Foru Aldundia/EHU, pp. 1-33.
- Alboiu, G., Barrie, M., and Frigeni, C. (2002) 'Derived reflexives in Romance: The case of SE', *Actes de l'ACL 2002/2002 CLA Proceedings*, pp. 1-14.
- Alcina Franch, J., and Blecua, J. M. (1975) *Gramática española*. Barcelona: Ariel.
- Alcover, A. M. (1916) 'Estudis de Sintaxis dialectal catalana', *Bolletí del diccionari de la llengua catalana*, IX, pp. 11-63.
- Alcover, A. M., and de B. Moll, F. (1929) *La flexió verbal en els dialectes catalans*. Barcelona: Anuari de l'Oficina Romànica de Lingüística i Literatura.

- Alencar, L. F., and Kelling, C. (2005) 'Are reflexive constructions transitive or intransitive? Evidence from German and Romance', in Butt, M., and Holloway King, T. (eds.) *Proceedings of the LFG05 Conference*, pp. 1-20. Available at: <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/10/lfg05-toc.html> (Accessed: 16th July 2016).
- Alexiadou, A., Anagnostopoulou, E., and Schäfer, F. (2006a) 'The properties of anticausatives crosslinguistically', in Frascarelli, M. (ed.) *Phases of Interpretation*. Berlin: Mouton de Gruyter, pp. 187-211.
- Alexiadou, A., Anagnostopoulou, E., and Schäfer, F. (2006b) 'The fine structure of (anti-)causatives', in Davis, C., Deal, A.-R., and Zabbal, Y. (eds.) *Proceedings of NELS 36*. Amherst: GLSA, pp. 115-128.
- Alexiadou, A., Anagnostopoulou, E., and Schäfer, F. (2015). *External arguments in transitivity alternations: a layering approach*. Oxford: Oxford University Press.
- Alibèrt, L. (1976) *Gramatica Occitana segun los parlars Lengadocians*. Montpelhièr: Centre d'Estudis Occitans.
- Alkire, T., and Rosen, C. G. (2010) *Romance languages: a historical introduction*. Cambridge/New York: Cambridge University Press.
- Alonso, A., and Henríquez Ureña, P. (1971) *Gramática castellana. Segundo curso*. Buenos Aires: Editorial Losada.
- Alsina, A. (1996) *The role of argument structure in grammar: Evidence from Romance*. Stanford: CSLI Publications.
- Alvar, M. (1953) *El dialecto aragonés [Biblioteca romanica hispanica]*. Madrid: Gredes.
- Alvar, M., and Pottier, B. (1983) *Morfología histórica del español*. Gredos.
- Álvarez Martínez, M. M. (1989) *El pronombre*. Madrid: Arco Libros.
- Anagnostopoulou, E. (2003) *The syntax of ditransitives: Evidence from clitics*. The Hague: Mouton.
- Anagnostopoulou, E. (2005) 'Strong and weak person restrictions: A feature checking analysis', in Heggie, L., and Ordóñez, P. (eds.) *Clitic and Affix Combinations: Theoretical Perspectives*. Amsterdam/Philadelphia, PA: John Benjamins, pp. 199-235.
- Anagnostopoulou, E. (2008) 'Notes on the Person Case Constraint in Germanic (with special reference to German)', in D'Alessandro, R., Fischer, S., and Hrafnbjarganson, H. (eds.) *Agreement Restrictions*. Berlin/New York: Mouton de Gruyter, pp. 15-47.
- Anderson, S. R. (1996) 'How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one', *The Linguistic Review*, 13, pp. 3-4.
- Anderson, S. R. (2005) *Aspects of the Theory of Clitics (Oxford Studies in Theoretical Linguistics)*. USA: Oxford University Press.
- Anderson, S. R. (2016) 'Romansh (Rumantsch)', in Ledgeway, A., and Maiden, M. (eds.) *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, pp.169-184.
- Arce, M. (1989) *Semantic structure and syntactic function: The case of Spanish se*. Doctoral Thesis. University of Colorado.

- Armstrong, G. (2013) 'Agentive Reflexive Clitics and Transitive Se Constructions in Spanish,' *Borealis: An International Journal of Hispanic Linguistics*, pp. 81-128.
- Arnold, J., Wasow, T., Asudeh, A., and Alrenga, P. (2004) 'Avoiding Attachment Ambiguities: the Role of Constituent Ordering', *Journal of Memory and Language*, 51, pp. 55-70.
- Arnold, J., Wasow, T., Losongco, A., and Ginstrom, R. (2000) 'Heaviness vs. Newness: The effects of complexity and information structure on constituent ordering', *Language*, 76, pp. 28-55.
- Arregi, K., and Nevins, A. (2007) 'Obliteration vs. Impoverishment in the Basque g-/z- Constraint', in *Proceedings of the 26th Penn Linguistics Colloquium, Penn Working Papers in Linguistics 13.1*. Philadelphia: Penn Linguistics Club, pp. 1-14.
- Aski, J., and Russi, C. (2010) 'The pragmatic functionality of atonic double object clitic clusters in 14th-century Florentine', *Folia Linguistica Historica*, 31, pp. 47-96.
- Astésano, C., and Bard, E. G. (2003) 'Structural and rhythmic influences on the occurrence of the initial accent in French', *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona*, pp. 503-506.
- Astésano, C., Bard, E. G., and Turk, A. (2002) 'Functions of the French initial accent: A preliminary study', in Bel, B., and Marlien, I. (eds.) *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence, pp. 139-142.
- Arsenijević, B. (2012) 'Evaluative reflections: Evaluative Dative Reflexive in South East Serbo-Croatian', in Fernández, B., and Exteparre, R. (eds.) *Variation in Datives*. Oxford: Oxford University Press, pp. 1-21.
- Auger, J. (1994) *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. Doctoral Thesis. University of Pennsylvania.
- Authier, J.-M., and Reed, L. A. (1992) 'On the syntactic status of French affected datives', *The Linguistic Review*, 9, pp. 295-311.
- Authier, J.-M., and Reed, L. A. (1996) 'On the Canadian French Middle', *Linguistic Inquiry*, 27.3, pp. 513-523.
- Avram, M. (2001) *Gramatica pentru toți*. Bucharest: Editura Humanitas.
- Ayres-Bennett, W. (2004) *Sociolinguistic variation in seventeenth-century France: methodology and case studies*. Cambridge/New York: Cambridge University Press.
- Azevedo, M. M. (1989) 'Vernacular Features in Educated Speech in Brazilian Portuguese', *Hispania*, 72.4, pp. 863-873.
- Baauw, S., and Delfitto, D. (2005) 'New views on reflexivity: Delay effects in Romance', *Probus*, 17.2, pp. 145-184.
- Babby, L. (1980) *Existential Sentences and Negation in Russian*. Ann Arbor, MI: Karoma Publishers.
- Babcock, S. S. (1970) *The Syntax of Spanish Reflexive Verbs*. The Hague: Mouton and Co.
- Bafile, L. (1993) *Fonologia prosodica e teoria metrica: accento, cliticizzazione e innalzamento vocalico in napoletano. Tesi per il conseguimento del titolo*. Thesis. Università degli studi di Firenze: Dipartimento di linguistica.
- Bafile, L. (1994) 'La riassegnazione postlessicale dell'accento nel napoletano', *Quaderni del 5 Dipartimento di Linguistica dell'università di Firenze*, 5, pp. 1-23.

- Baker, M., Johnson, K., and Roberts, I. (1989) 'Passive arguments raised', *Linguistic Inquiry*, 20, pp. 219-252.
- Barlow, M. (2000) 'Usage, Blends, and Grammar', in Barlow, M., and Kemmer, S. (eds.) *Usage-based models of language*. Stanford: CSLI, pp. 315-345.
- Baron, I., Herslund, M., and Sørensen, F. (2001) *Dimensions of possession*. Amsterdam/Philadelphia: John Benjamins Pub. Co. Available at: <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622462> (Accessed: 21st October 2012).
- Barrenechea, A. M., and Manacorda de Rosetti, M. (1977) 'La voz pasiva en el español hablado en Buenos Aires', in Lope Blanch, J. M. (ed.) *Estudios sobre el español hablado en las principales ciudades de América*. México: Universidad Nacional Autónoma de México, pp. 395-405.
- Barry, A. K. (1987) 'Clitic Pronoun Position in Thirteenth-Century Spanish', *Hispanic Review*, 55.2, pp. 213-220.
- Bastida, S. (1976) 'Restricciones de orden en las secuencias de clíticos del castellano: dos requisitos', in Sánchez de Zavala, V. (ed.) *Estudios de gramática generativa*. Barcelona: Labor, pp. 59-99.
- Battaglia, S., and Pernicone, V. (1968) *Grammatica italiana*. Torino: Loescher.
- Bayle, L. (1989) *Grammaire provençale: Avec exercices, vocabulaires, textes de lecture, tableaux de la conjugaison [Estudi prouvençau]*. Berre l'Étang: l'Astrado.
- Beaver, D., Francez, I., and Levinson, D. (2006) 'Bad subject: (Non-)canonicity and NP distribution in existentials', in Howell, E. G. and Howell, J. (eds.) *Proceedings of Semantic and Linguistic Theory XV*. Ithaca, NY: Cornell University Press, pp. 19-43.
- Beavers, J., and Koontz-Garboden, A. (2013a) 'In defense of the reflexivization analysis of anticausativization', *Lingua*, 131, pp. 199-216.
- Beavers, J., and Koontz-Garboden, A. (2013b) 'Complications in diagnosing lexical meaning: A rejoinder To Horvath & Siloni (2013)', *Lingua*, 134, pp. 210-218.
- Béjar, S. (2003) *Phi-syntax: A theory of agreement*. Doctoral Thesis. University of Toronto.
- Béjar, S., and Řezáč, M. (2003) 'Person licensing and the derivation of PCC effects', in Pérez-Leroux, A.-T., and Roberge, Y. (eds.) *Romance linguistics: Theory and acquisition*. Amsterdam: John Benjamins, pp. 49-61.
- Béjar, S., and Řezáč, M. (2009) 'Cyclic agree', *Linguistic Inquiry*, 40, pp. 35-73.
- Belletti, A. (2001) 'Inversion as focalization', in Hulk, A., and Pollock, J.-Y. (eds.) *Subject inversion in Romance and the theory of universal grammar*. New York: Oxford University Press, pp. 60-90.
- Belletti, A., and Rizzi, L. (1981) 'The Syntax of ne: Some Theoretical Implications', *The Linguistic Review*, 1, pp. 117-154.
- Belletti, A., and Rizzi, L. (1988) 'Psych-verbs and Theta-theory', *Natural Language and Linguistic Theory*, 6, pp. 291-352.
- Bello, A. (1847) *Gramática de la lengua castellana, con notas de R.J. Cuervo y N. Alcalá Zamora*. Buenos Aires: Sopena.
- Bello, A., and Cuervo, R. (1960) *Gramática de la lengua Castellana*. 6th edition. Buenos Aires: Sopena.

- Benincà, P. (1988) 'Fenomenologia della dislocazione a sinistra', in Renzi, L. (ed.) *Grande grammatica italiana di consultazione. La frase. I sintagmi nominale e preposizionale, vol. I*. Bologna: Mulino, pp. 153-194.
- Benincà, P. (1999) 'Between morphology and syntax: On the verbal morphology of some Alpine dialects', in Mereu, L. (ed.) *Boundaries of Morphology and Syntax*. (Current Issues in Linguistic Theory 180). Amsterdam/Philadelphia: John Benjamins, pp. 11-30.
- Benincà, P. (2007) 'Clitici e ausiliari: gh ò, z é', in Bentley, D. and Ledgeway, A. (eds.) *Sui dialetti italo-romanzi. Saggi in onore di Nigel Vincent*. (The Italianist 27 (Special Supplement 1)), pp. 27-47.
- Benincà, P., and Cinque, G. (1993) 'Su alcune differenze tra enclisi e proclisi', in Folena, G. (ed.) *Omaggio a Gianfranco Folena*. Padova: Editoriale Programma, pp. 2313-2326.
- Benincà, P., and Poletto, C. (2004) 'Topic, focus, and V2: Defining the CP sublayers', in Rizzi, L. (ed.) *The Structure of CP and IP: The Cartography of Syntactic Structures*. New York: Oxford University Press, pp. 52-75.
- Benincà, P., and Poletto, C. (2005) 'The Third Dimension of Person Features', in Cornips, L., and Corrigan, K. P. (eds.) *Syntax and Variation. Reconciling the Biological and the Social*, pp. 265-299.
- Benincà, P., and Tortora, C. (2009) 'Towards a finer-grained theory of Italian participial clausal architecture', *Working Papers in Linguistics. University of Pennsylvania*, 15, pp. 16-26.
- Benincà, P., Salvi, G., and Frison, L. (2001) 'L'ordine degli elementi della frase e le costruzioni marcate', in Renzi, L., Salvi, G., and Cardinaletti, A. (eds.) *Grande grammatica italiana di consultazione*. Bologna: Il Mulino, pp. 129-239.
- Bentley, D. (2004) 'Ne-Cliticisation and Split Intransitivity', *Journal of Linguistics*, 40.2, pp. 219-262.
- Bentley, D. (2010) 'Principles of subject markedness in Romance', *Archivio Glottologico Italiano*, 95, pp. 152-189.
- Bentley, D. (2011) 'Sui costrutti esistenziali sardi. Effetti di definitezza, deissi, evidenzialità', *Zeitschrift für Romanische Philologie*, 127.1, pp. 111-140.
- Bentley, D., Ciconte, F. M., and Cruschina, S. (2013) 'Micro-variation in subject agreement: The case of existential pivots with split focus in Romance', *Rivista di Linguistica*, 25.1, pp. 15-43.
- Benveniste, E. (1966a) 'Être et avoir dans leur fonction linguistique', in Benveniste, E. (ed.) *Problèmes de linguistique générale*. Paris: Gallimard, Chapter XVI.
- Benveniste, E. (ed.) (1966b) *Problèmes de linguistique générale*. Paris: Gallimard.
- Berman, R. (1981) 'Dative marking of the affectee role: data from Modern Hebrew', *MIT Working Paper in Linguistics*, 3, pp. 35-59.
- Berretta, M. (1985a) 'I pronomi clitici nell'italiano parlato', in Holtus, G., and Radtke, E. (eds.) *Gesprochenes Italienisch in Geschichte und Gegenwart*. Tübingen: Narr., pp. 185-224.
- Berretta, M. (1985b) 'Ci vs. gli: un microsistema in crisi?', in Franchi de Bellis, A., and Savoia, L. M. (eds.) *Sintassi e morfologia della lingua italiana d'uso: teorie e applicazioni descrittive. Atti del XVII Congresso SLI, Urbino, September 11 - 13, 1983*. Roma: Bulzoni, pp. 117-133.

- Berretta, M. (1989) 'Tracce di coniugazione oggettiva in italiano', in Foresti, F., Rizzi, E., and Benedini, P. (eds.) *L'italiano tra le lingue romanze. Atti del XX Congresso della Società Linguistica Italiana. Bologna 25–27 settembre 1986*. Roma: Bulzoni, pp. 125-150.
- Berruto, G. (1985a) 'Discocazioni a sinistra e "grammatica" dell'italiano parlato', in De Bellis, A. F., and Savoia, L. M. (eds.) *Sintassi e morfologia della lingua italiana d'uso. Teorie e applicazioni descrittive. Atti del XVII Congresso SLI, Urbino, September 11–13, 1983*. Roma: Bulzoni, pp. 79-82.
- Berruto, G. (1985b) 'Per una caratterizzazione del parlato: l'italiano parlato ha un'altra grammatica?', in Holtus, G., and Radtke, E. (eds.) *Gesprochenes Italienisch in Geschichte und Gegenwart*. Tübingen: Narr., pp. 120-153.
- Berruto, G. (1986) 'Le dislocazioni a destra in italiano', in Stammerjohann, H. (ed.) *Tema/Rema in italiano*. Tübingen: Narr., pp. 55-69.
- Berruto, G. (1987) *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci.
- Bhatt, R., and Pancheva, R. (2006) 'Implicit Arguments', in Everaert, M., and van Riemsdijk, H. (eds.) *The Blackwell Companion to Syntax*, Vol. II. Malden, Mass: Blackwell, pp. 558-588.
- Bianchi, V. (2006) 'On the syntax of personal arguments', *Lingua*, 116.12, pp. 2023-2067.
- Birabent, J.-P., and Salles-Loustau, J. (1989) *Memento grammatical du gascon*. Pau: Centre Régional de Documentation Pédagogique.
- Bisang, W., Himmelmann, N., and Wiemer, B. (eds.) (2004) *What Makes Grammaticalization: A Look at Its Fringes and Its Components*. (Trends in Linguistics. Studies and Monographs 158). Berlin/New York: Mouton de Gruyter.
- Blasco Ferrer, E. (1986) *La lingua sarda contemporanea: Grammatica del logudorese e del campidanese [Sintesi storica]*. Cagliari: Edizioni della Torre.
- Blasco Ferrer, E. (1988) *Le parlate dell'Alta Ogliastra*. Cagliari: Edizioni della Torre.
- Blasco Ferrer, E. (2003) 'Tipologia delle presentative romanze e morfosintassi storica. Fr. C'est e prov. –i (estai, fai, plai)', *Zeitschrift für romanische Philologie*, 119.1, pp. 51-90.
- Bock, J. K. (1982) 'Toward a Cognitive Psychology of Syntax: Information Processing Contributions to Sentence Formulation', *Psychological Review*, 89, pp. 1-47.
- Boland, J. E. (1997) 'Resolving Syntactic Category Ambiguities in Discourse Context: Probabilistic and Discourse Constraints', *Journal of Memory and Language*, 36, pp. 588-615.
- Bolinger, D. (1969) 'Of undetermined Nouns and indeterminate Reflexives', *Romance Philology*, 22.4, pp. 484-489.
- Bolognesi, R. (1998) *The phonology of Campidanian Sardinian: A unitary account of a self-organizing structure*. Dordrecht: ICG Printing.
- Boneh, N., and Nash, L. (2011) 'High and Higher Applicatives: The Case of French Non-core Datives', in Byram Washburn, M., McKinney-Bock, K., Varis, E., Sawyer, A., and Tomaszewicz, B. (eds.) *Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadia Proceedings Project, pp. 60-68.

- Boneh, N., and Sichel, I. (2010) *Deconstructing Possession*. Masters Thesis. Linguistics Department. The Hebrew University of Jerusalem.
- Bonet, E. (1991) *Morphology after Syntax: Pronominal clitics in Romance languages*. Doctoral Thesis. MIT.
- Bonet, E. (1993) '3rd Person Pronominal Clitics In Dialects Of Catalan', *Catalan Working Papers in Linguistics*, 3.1, pp. 85-111.
- Bonet, E. (1994) 'The Person-Case Constraint: A Morphological Approach', *Morphology-Syntax Connection number 22 in MIT Working Papers in Linguistics*, 22, pp. 33-52.
- Bonet, E. (1995a) 'Feature Structure of Romance Clitics', *Natural Language & Linguistic Theory*, 13.4, pp. 607-647.
- Bonet, E. (1995b) 'The Where and How of Clitic Order', *Revue Québécoise De Linguistique*, 24.1, pp. 61-81.
- Bonet, E. (2002) 'Cliticització', in Solà, J., Lloret, M. R., Mascaró, J., and Pérez Saldanya, M. (eds.) *Gramàtica del Català Contemporani*. Barcelona: Empúries, pp. 933-989.
- Bonet, E. (2008) *The Person-Case Constraint and Repair Strategies*. Berlin, Germany: Mouton de Gruyter.
- Bonet, E., and Torres-Tamarit, F. (2010) *Les Formes D'imperatiu Seguides De Clitic: Un Cas De Conservadorisme Lèxic*. Esborrany. Available at: <http://filcat.uab.cat/clt/publicacions/reports/pdf/GGT-10-04.pdf> (Accessed: 1st January 2016).
- Bonnaud, P. (1992) *Grammaire générale de l'auvergnat à l'usage des arvernaisants*. Chamalières: Cercle Terre d'Auvergne.
- Borer, H., and Grodzinsky, Y. (1986) 'Syntactic Cliticization: The Case of Hebrew Dative Clitics', in Borer, H. (ed.) *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. New York: Academic Press, pp. 175-218.
- Borer, H. (2005) *The Normal Course of Events. Structuring Sense, Volume II*. Oxford, Oxford University.
- Bosse, S., and Bruening, B. (2011) 'Benefactive Versus Experiencer Datives', in Washburn, M. B., McKinney-Bock, K., Varis, E., Sawyer, A., and Tomaszewicz, B. (eds.) *Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadia Proceedings Project, pp. 69-77. Available at: www.lingref.com. (Accessed: 1st January 2016).
- Boucher, C. (2013) 'Double Object Clitic Ordering in a Dialect of Corsican'. Available at: <https://linguistics.stonybrook.edu/sites/default/files/uploads/Boucher.pdf> (Accessed: 1st January 2016).
- Bowerman, M. (1988) 'The "No Negative Evidence" Problem: How Do Children Avoid Constructing an Overly General Grammar?', in Hawkins, J. A. (ed.) *Explaining Language Universals*. Oxford: Blackwell, pp. 73-101.
- Bowerman, M. (1996) 'Argument Structure and Learnability: is a Solution in Sight?', in Johnson, J., Juge, M. L., and Moxley, J. L. (eds.) *Parasession on Learnability and grammatical Theory*. Berkeley Linguistics Society 22, pp. 454-468.
- Bowerman, M., and Choi, S. (2003) 'Space under Construction: Language-Specific Spatial Categorization in First Language Acquisition', in Gentner, D., and Goldin-Meadow, S. (eds.) *Language in Mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press, pp. 387-427.
- Bowers, J. (1993) 'The syntax of predication', *Linguistic Inquiry*, 24, pp. 591-656.

- Braine, M. D. S., and Brooks, P. J. (1995) 'Verb Argument Structure and the Problem of Avoiding an Overgeneral Grammar', in Tomasello, M., and Merriman, W. E. (eds.) *Beyond Names for Things: Young Children's Acquisition of Verbs*. Hillsdale, New York: Lawrence Erlbaum Assoc., pp. 353-376.
- Brambilla Ageno, F. (1964) *Il verbo nell'italiano antico*. Milano/Napoli: Ricciardi.
- Bransford, J. D., and Franks, J. J. (1972) 'The abstraction of linguistic ideas: A review', *Cognition*, 1, pp. 211-249.
- Bransford, J. D., and McCarrell, N.S. (1977) 'A sketch of a cognitive approach to comprehension: some thoughts about understanding what it means to comprehend', in Johnson-Laird, P. N., and Wason, P. C. (eds.) *Thinking, Readings in Cognitive Science*. Cambridge: Cambridge University Press, pp. 377-399.
- Bresnan, J. (2006) 'Is syntactic knowledge probabilistic? Experiments with the English dative alternation', *Paper presented at the International Conference on Linguistic Evidence*. Available at: <https://web.stanford.edu/~bresnan/download.html> (Accessed: 18th July 2016).
- Bresnan, J., and Kanerva, J. M. (1989) 'Locative inversion in Chichewa: A case study of factorization in grammar', *Linguistic Inquiry*, 20.1, pp. 1-50.
- Bresnan, J., and Moshi, L. (1990) 'Object asymmetries in comparative Bantu syntax', *Linguistic Inquiry*, 21, pp. 147-185.
- Brinton, L., and Traugott, E. C. (2005) *Lexicalization and Grammaticalization in Language Change*. Cambridge: Cambridge University Press.
- Britt, M. A., Perfetti, C. A., Garrod, S., and Rayner, K. (1992) 'Parsing in Discourse: Context Effects and Their Limits', *Journal of Memory and Language*, 31, pp. 293-314.
- Broe, M. B., and Pierrehumbert, J. B. (2000) 'Introduction', in Broe, M. B., and Pierrehumbert, J. B. (eds.) *Acquisition and the Lexicon*. Cambridge: Cambridge University Press, pp. 1-7.
- Brousseau, A-M., and Ritter, E. (1991) 'A non-unified analysis of agentive verbs', in Bates, D. (ed.) *Proceedings of the 10th West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI, pp. 53-64.
- Browman, C. P., and Goldstein, L. (1992) 'Articulatory phonology: an overview', *Phonetica*, 49, pp. 155-80.
- Bruening, B. (2006) *The morphosyntax and semantics of verbal reciprocals*. University of Delaware.
- Bruhn de Garavito, J., Lamarche, J., and Heap, D. (2002) 'French and Spanish se: Underspecified, not reflexive', in Somesfalean, S. (ed.) *Proceedings of the 2002 Annual meeting of the Canadian Linguistics Association*. Montréal, Quebec: UQÀM, pp. 42-54.
- Brunot, F. (1965). *La pensée et la langue*. Paris: Masson
- Brunot, F. (1966) *Histoire de la langue française des origines à nos jours, tome 3, La formation de la langue classique 1600 – 1660*. Paris: Colin.
- Brusewitz, V. (1905) *Étude historique sur la syntaxe des pronoms personnels dans la langue des félibres*. Thesis. Upsal.
- Bull, W. (1965) *Spanish for teachers: applied linguistics*. Nueva York: Ronald Press Company.
- Burston, J. L. (1983) 'Clitic Object sequence and Cooccurrence Restrictions in French', *Linguistic Analysis*, 11.3, pp. 247-275.

- Burzio, L. (1986) *Italian syntax: a government-binding approach*. Dordrecht: Kluwer Academic Publishers.
- Butt, J., and Benjamin, C. (1994) *A New Reference Grammar of Modern Spanish*. London/New York: Edward Arnold.
- Bybee, J. L. (2001) *Phonology and language use*. Cambridge/New York: Cambridge University Press.
- Bybee, J. L., and Hopper, P. J. (2001) *Frequency and the emergence of linguistic structure*. Amsterdam/Great Britain: J. Benjamins. Available at: <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622579> (Accessed: 21st October 2012).
- Bybee, J. L., Perkins, R. D., and Pagliuca, W. (1994) *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Bybee, J., and Eddington, D. (2006) 'A usage-based approach to Spanish verbs of "becoming"', *Language*, 82, pp. 323-355.
- Bybee, J., and Pagliuca, W. (1987) 'The evolution of future meaning', in Ramat, A. G., Carruba, O., and Bernini, L. (eds.) *Papers from the 7th International Conference on Historical Linguistics*. (Current Issues in Linguistic Theory 48). Amsterdam/Philadelphia: John Benjamins, pp. 108-122.
- Cable, S. (2014) 'Reflexives, Reciprocals and Contrast', *Journal of Semantics*, 31.1, pp. 1-41.
- Cabr  i Monet, T. (1994) 'Minimality in the Catalan truncation process', *Catalan Working Papers in Linguistics*, 4.1, pp. 1-21.
- Cabredo Hofherr, P. (2006) "'Arbitrary" pro and the theory of pro-drop', in Ackema, P., Brandt, P., Schoorlemmer, M., and Weerman, F. (eds.) *Agreement and Arguments*. Oxford: Oxford University Press, pp. 230-258.
- Campanini, C. and Sch fer, F. (2011) *Optional Se-Constructions in Romance: Syntactic Encoding of Conceptual Information*. Universit t Stuttgart.
- Campmany, E. (2008) *Difer ncies fonol giques entre diversos estils de parla al catal  central septentrional*. Doctoral Thesis. University of Barcelona.
- Campos, H. (1999) 'Transitividad e intransitividad', in Bosque, I., and Demonte, V. (eds.) *Gram tica descriptiva de la lengua espa ola*. Madrid: Espasa Calpe, pp. 519-74.
- Camproux, C. (1958) *Etude syntaxique des parlers g vaudonais [Th se d'Etat]*. Paris: Presses Universitaires de France.
- Cardinaletti, A. (2008) 'On different types of clitic clusters', in De Cat, C., and Demuth, K. (eds.) *The Bantu-Romance connection a comparative investigation of verbal agreement, DPs, and information structure*. Amsterdam/Philadelphia: John Benjamins, pp. 41-82. Available at: <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622262> (Accessed: 22nd October 2012).
- Cardinaletti, A., and Egerland, V. (2010) 'I pronomi personali e riflessivi', in Salvi, G., and Renzi, L. (eds.) *Grammatica dell'italiano antico*. Bologna: Il Mulino, pp. 414-450.
- Cardinaletti, A., and Giusti, G. (2006) 'The syntax of quantified phrases and quantitative clitics', in Everaert, M., and van Riemsdijk, H. (eds.) *The Blackwell Companion to Syntax*. Oxford: Blackwell, pp. 23-93.
- Cardinaletti, A., and Repetti, L. (2008) 'The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects', *Linguistic Inquiry*, 39.4, pp. 523-563.

- Cardinaletti, A., and Repetti, L. (2009) 'Phrase-level and word-level syllables: Resyllabification and prosodization of clitics', in Grijzenhout, J., and Kabak, B. (eds.) *Phonological Domains: Universals and Derivations*. Berlin: Mouton de Gruyter, pp. 79-104.
- Cardinaletti, A., and Starke, M., (1994) 'The typology of structural deficiency. On the three grammatical classes', *University of Venice Working Paper in Linguistics*, 4, pp. 41-109.
- Cardinaletti, A., and Starke, M. (1999) 'The typology of structural deficiency: A case study of the three classes of pronouns', *Empirical Approaches to Language Typology*, pp. 145-234.
- Carlier, A. (2007) 'From Preposition to Article : The Grammaticalization of the French Partitive', *Studies in Language*, 31, pp. 1-49.
- Carratalá, E. (1980) *Morfosintaxis del castellano actual*. Barcelona: Labor.
- Cartagena, N. (1972) *Sentido y estructura de las construcciones pronominales en español*. Univ. de Concepción: Publicaciones del Instituto Central de Lenguas.
- Casanova Herrero, E. (1990) 'Aproximació a una gramàtica contrastiva dels dialectes catalans al segle XVIII: la combinació binària de pronoms personals febles de 3a. persona', *Actes del Vuitè Col·loqui Internacional de Llengua i Literatura Catalanes*, pp. 57-80.
- Cattaneo, A. (2009) *It Is All About Clitics: The Case of a Northern Italian Dialect Like Bellinzonese*. Doctoral Thesis. New York University.
- Cennamo, M. (1999) 'Inaccusatività tardo-latina e i suoi riflessi in testi italiani antichi centro-meridionali', *Zeitschrift für romanische Philologie*, 115.2, pp. 300-331.
- Cennamo, M. (2000) 'Patterns of "active" syntax in Late Latin pleonastic reflexives', in Smith, J. C., and Bentley, D. (eds.) *Historical Linguistics*. (Current Issues in Linguistic Theory 161). Amsterdam/Philadelphia: John Benjamins, pp. 35-55.
- Cennamo, M. (2011) 'Impersonal constructions and accusative subjects in Late Latin', in Malchukov, A., and Siewierska, A. (eds.) *Impersonal Constructions: a Cross-linguistic Perspective*. Amsterdam: Benjamins, pp. 169-188.
- Cennamo, M., and Sorace, A. (2007) 'Unaccusativity at the syntax-lexicon interface: evidence from Paduan', in Aranovich, R. (ed.) *Split Auxiliary Systems. A Cross-linguistic Perspective*. Amsterdam: John Benjamins, pp. 65-100.
- Centineo, G. (1995) 'The Distribution of si in Italian Transitive/Inchoative Pairs', in Mandy, M., and Galloway, T. (eds.) *Proceedings from Semantics and Linguistic Theory V*. Ithaca, NY: CLC Publications, Cornell University, pp. 54-71.
- Chabaneau, C. (1876[1980]) *Grammaire Limousine*. (Reprinted from *Revue des langues romanes*, 2-9. Paris: Maisonneuve, 1871-1876). Marseille: Laffitte Reprints.
- Chappell, H., and McGregor, W. (eds.) (1996) *The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part-Whole Relation*. Berlin: Mouton de Gruyter.
- Charaud, J. (2000) 'Les variétés régionales du français', in Antoine, G., and Cerquiglini, B. (eds.) *Histoire de la langue française 1945-2000*. Paris: CNRS, pp. 643-64.
- Chereches, A. (2014) 'A Prosodic Analysis of Romanian Pronominal Clitics', *U. Penn Working Papers in Linguistics*, 20.1, pp. 51-60.

- Chierchia, G. (2004) 'A semantics for unaccusatives and its syntactic consequences', in Alexiadou, A., Anagnostopoulou, E., and Everaert, M. (eds.) *The unaccusativity puzzle: explorations of the syntax-lexicon interface*. Oxford: Oxford University Press, pp. 22-59.
- Chitoran, I. (2002) *The Phonology of Romanian: A Constraint-Based Approach*. Walter de Gruyter.
- Chomsky, N. (1957) *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1981) *Lecture on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1995) *The minimalist program*. Cambridge, Mass.: The MIT Press.
- Chomsky, N. (2000) 'Minimalist Inquiries: The Framework', in Martin, R., Michaels, D., and Uriagereka, J. (eds.) *Step by Step: Essays of a Minimalist Syntax in Honor of Howard Lasnik*. Cambridge: MIT Press, pp. 89-155.
- Chusé, L. A., and Chuan-chusé Lagraba, M. (1987) *De la gramatica de lo cheso, fabla alto-aragonesa*. Concello de la villa de Val d'Echo: Grupo d'Estudios de la Fabla Chesa.
- Ciconte, F. M. (2009) 'Pro-forms in existential constructions of early Italo-Romance vernaculars', in Kaiser, G. A., and Remberger, E.-M. (eds.) *Proceedings of the Workshop "Null-subjects, expletives, and locatives in Romance"*. (Arbeitspapier 123). Fachbereich Sprachwissenschaft: Universität Konstanz, pp. 183-98.
- Ciconte, F. M. (2010) *Existential constructions of Early Italo-Romance Vernaculars (To appear in the Monograph Series of The Philological Society)*. Doctoral Thesis. University of Manchester.
- Cinque, G. (1988) 'On Si Constructions and the Theory of Arb', *Linguistic Inquiry*, 19, pp. 521-582.
- Cinque, G. (1995) *Italian syntax and universal grammar*. Cambridge: Cambridge University Press.
- Ciucivara, O. (2009) *A Syntactic Analysis of Pronominal Clitic Clusters in Romance. The view from Romanian*. Doctoral Thesis. New York University.
- Clark, E. V., and Carpenter, K. L. (1989) 'On children's uses of from, by, and with in oblique noun phrases', *Journal of Child Language*, 16, pp. 349-364.
- Colantoni, L., Cuervo, M. C., and Hualde, J. I. (2010) 'Stress as a symptom', *Paper presented at LSRL 40, University of Washington, March 26-28, 2010*.
- Colomina i Castanyer, J. (1985) 'Alguns casos de variació morfològica en valenci meridional alacantí', in Gil-Albert, J. (ed.) *L'Alacantí, un estudi sobre la variació lingüística*. Alacant: Institut d'Estudis.
- Colomina i Castanyer, J. (1991) *El Valencià de la Marina Baixa. Estudis del Valencià Actual*. València: Generalitat Valenciana.
- Company, C. (1998) 'The interplay between form and meaning in language change. Grammaticalization of cannibalistic datives in Spanish', *Studies in Language*, 22.3, pp. 529-565.
- Company, C. (2001) 'Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language', *Studies in Language*, 25.1, pp. 1-47.
- Comrie, B. (1981) *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

- Conroy, A. M. (2007) 'The Personal Dative in Appalachian English as a Reflexive Pronoun', in Omaki, A., Ortega-Santos, I., Sprouse, J., and Wagers, M. (eds.) *University of Maryland Working Papers in Linguistics 16*. College Park, MD: UMWPiL, pp. 63-88.
- Contreras, L. (1964) 'Significados y funciones del "se"', *Onomázein*, 9, pp. 95-104.
- Corblin, F. (1995) *Les formes de reprise dans le discours anaphores et chaînes de référence*. Rennes: Presses universitaires de Rennes.
- Cordin, P. (2001) 'I pronomi riflessivi', in Renzi, L., Salvi, G., and Cardinaletti, A. (eds.) *Grande grammatica italiana di consultazione. Volume I*. Bologna: Il Mulino, pp. 600-602.
- Cordin, P., and Calabrese, A. (2001) 'I pronomi personali', in Renzi, L., Salvi, G., and Cardinaletti, A. (eds.) *Grande grammatica italiana di consultazione*. Bologna: Il Mulino, pp. 549-606.
- Cornilescu, A. (2009) 'Restructuring strategies of the Romanian a fi "be"', in Kaiser, G. A., and Remberger, E.-M. (eds.) *Proceedings of the workshop "Null-subjects, expletives and locatives in Romance"*. Universität Konstanz: Fachbereich Sprachwissenschaft, pp. 199-230.
- Cortelazzo, M., and Zolli, P. (1999) *Il Nuovo Etimologico*. Bologna: Zanichelli.
- Cortés, C., and Gavarró, A. (1997) 'Subject-Object Assymetries and the Clitic En', in Back, J., and Motapanyane, V. (eds.) *Clitics, Pronouns and Movement*. Amsterdam: John Benjamins, pp. 39-62.
- Côté, M.-H. (2014) 'Liaison et assibilation en français laurentien', in Soum-Favaro, C., Coquillon, A., and Chevrot, J.-P. (eds.) *La liaison : approches contemporaines*. Berne: Peter Lang, pp. 9-32.
- Couquaux, D. (1978) 'Sur une incompatibilité de pronoms clitiques en français', *Linguisticae Investigationes*, 2, pp. 211-214.
- Craig, C. (1976) 'Properties of Basic and Derived Subjects in Jacalteco', in Li, C. N. (ed.) *Subject and Topic*. New York: Academic Press, pp. 99-123.
- Croft, W. (2000) *Explaining language change*. London: Longman.
- Croft, W., and Cruse, D. A. (2004) *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Cruschina, S. (2012) 'Focus in Existential Sentences', in Bianchi, V., and Chesi, C. (eds.) *Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday*. Siena: CISCL Press, pp. 77-107. Available at: <http://www.ciscl.unisi.it/gg60/papers/cruschina.pdf> (Accessed: 1st January 2016).
- Cruschina, S. (2014) 'Existential and locative constructions in Italo-Romance', *L'Italia dialettale*, 75, pp. 55-80.
- Cuervo, M. C. (2003) *Datives at large*. Thesis. Massachusetts Institute of Technology. Available at: <http://dspace.mit.edu/handle/1721.1/7991> (Accessed: 6th November 2014).
- Cuervo, M. C. (2008) 'La alternancia causativa y su interacción con argumentos dativos', *RLA. Revista de Lingüística Teórica y Aplicada*, 46.1, pp. 55-79.
- Cuervo, M. C. (2010) 'Against ditransitivity', *Probus*, 22, pp. 151-180.
- Culicover, P. W. (1999) *Syntactic Nuts*. Oxford: Oxford University Press.
- Culicover, P. W., and Nowak, A. (2003) *Dynamical Grammar*. Oxford: Oxford University Press.

- D'Achille, P. (1990) *Sintassi del parlato e tradizione scritta della lingua italiana*. Roma: Bonacci.
- D'Alessandro, R. (2004) *Impersonal si constructions. Agreement and interpretation*. Doctoral Thesis. Universität Stuttgart.
- Dardano, M., and Trifone, P. (1995) *La nuova grammatica della lingua italiana*. Bologna: Zanichelli.
- Dauzat, A. (1927) *Les patois*. Paris.
- Davis, H., and Demirdache, H. (2000) 'On Lexical Verb Meanings: Evidence from Salish', in Tenny, C., and Pustejovsky, J. (eds.) *Events as Grammatical Objects*. Stanford, CA: CSLI, pp. 97-142.
- de Benito Moreno, C. (2015) 'Pero se escondíamos como las ratas: syncretism in the reflexive paradigm in Spanish and Catalan', *Isogloss*, pp. 95-127.
- de Borja Moll, F. (1968) *Gramàtica Catalana*. Palma de Mallorca: Editorial Moll.
- de Borja Moll, F. (1980) *El parlar de Mallorca, referida especialment a les illes Balears*. Barcelona: Editorial Barcino.
- De Cuyper, G. (2006) *La estructura léxica de la resultatividad y su expresión en las lenguas germánicas y románicas*. Munich: Lincom.
- de Fourvières, X. (1986) *Grammaire et guide de la conversation provençales*. Raphèle-les-Arles: Culture provençale et méridionale.
- De Mauro, T. (1999-2000) *GRADIT: Grande dizionario italiano dell'uso*. Torino: UTET.
- De Mello, G. (1978) 'On the use of 'por' plus agent with 'se' constructions', *Hispanica*, 61, pp. 323-327.
- De Mello, G. (1992) 'Se los for se lo in the spoken cultured Spanish of eleven cities', *Hispanic Journal*, 13.1, pp. 165-179.
- De Mello, G. (1995) 'Concordancia entre el verbo pronominal de tercera persona y su sustantivo: se venden flores vs. se vende flores', *Anuario de letras XXXIII*, pp.58-82.
- de Miguel, E. (1999) 'El aspecto léxico', in Bosque, I., and Demonte, V. (eds.) *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe, pp. 2977-3060.
- de Miguel, E., and Fernández Lagunilla, M. (2000) 'El operador aspectual se', *Revista Española de Lingüística*, 30.1, pp. 13-43.
- De Molina Redondo, J. A. (1974) *Usos de "se": Cuestiones sintácticas y léxicas*. Madrid: Sociedad General Española de Librería, S.A.
- de Swart, H. (1998) 'Aspect shift and coercion', *Natural Language and Linguistic Theory*, 16, pp. 347-85.
- Deacon, T. W. (2003) 'Universal Grammar and Semiotic Constraints', in Christiansen, M. H., and Kirby, S. (eds.) *Language evolution*. Oxford: Oxford University Press, pp. 111-139.
- Deane, P. (1987) 'English Possessives, Topicality, and the Silverstein Hierarchy', *Berkeley Linguistics Society*, 13, pp. 65-76.

- Déchainé, R.-M., and Wiltschko, M. (2004) 'Deconstructing reciprocals', *Paper presented at the western conference on linguistics*. University of Southern California.
- Delais-Roussarie, E. (1995) *Pour une approche parallèle de la structure prosodique : étude de l'organisation prosodique et rythmique de la phrase française*. Doctoral Thesis. Université de Toulouse.
- Delais-Roussarie, E. (1999) 'Accentuation et réalisation des clitiques en français', *Cahiers de Grammaire*, 24, pp. 17-37.
- Delbecque, N. (1995) 'Towards a Cognitive Account of the Use of the Prepositions *por* and *para* in Spanish', in Casad, E. H. (ed.) *Cognitive Linguistics in the Redwoods. The expansion of a New Paradigm in Linguistics*. Berlin: Mouton de Gruyter.
- Delbecque, N. and Lamiroy, B. (1996) 'Towards a typology of the Spanish dative', in van Belle, W., and van Langendonck, W. (eds.) *Case and grammatical relations across languages: The Dative*. Amsterdam: John Benjamins, pp. 73-117.
- Delfitto, D. (2002) 'On the semantics of pronominal clitics and some of its consequences', *Catalan Journal of Linguistics*, 1, pp. 41-69.
- Demonte, V. (1994) *Gramática del español*. México: Colegio de México.
- Demonte, V. (1995) 'Dative alternation in Spanish', *Probus*, 7, pp. 5-30.
- Di Cristo, A. (1999) 'Vers une modélisation de l'accentuation du français: première partie', *French Language Studies*, 9, pp. 143-179.
- Di Cristo, A. (2000) 'Vers une modélisation de l'accentuation du français: seconde partie', *French Language Studies*, 10, pp. 27-44.
- Di Tullio, A. (1997) *Manual de gramática del español : desarrollos teóricos, ejercicios, soluciones*. Buenos Aires: Edicial.
- Diaconescu, C. R. (2004) *Romanian Applicative Constructions*. Doctoral Thesis. University of Ottawa.
- Diesing, M. (1992) *Indefinites*. Cambridge: MIT Press.
- Dik, S. (1989) *The Theory of Functional Grammar. Part I: The structure of the clause*. Dordrecht: Foris Publications.
- Dinnsen, D. (1972) 'Additional constraints on clitic order in Spanish', in Saciuk, B., and Casagrande, J. (eds.) *Generative Studies in Romance Languages*. Rowley, MA: Newbury House, pp. 175-183.
- D'Introno, F., González, V., and Rivas, J. (2007) 'Aspectos sintácticos y semánticos del pronombre *se*', *Boletín de Lingüística*, 28, pp. 5-25.
- Dobrovie-Sorin, C. (1994a) *The Syntax of Romanian: Comparative Studies in Romance*. (Series Studies in generative Grammar 40). Berlin, New York: Mouton de Gruyter.
- Dobrovie-Sorin, C. (1994b) 'CD, wh-movement, and quantification in Romanian', *Linguistic Inquiry*, 21, pp. 351-391.
- Dobrovie-Sorin, C. (1998) 'Impersonal *se*-constructions in Romance and the passivization of unergatives', *Linguistic Inquiry*, 29, pp. 399-437.

- Dobrovie-Sorin, C. (2006) 'The SE-Anaphor and its Role in Argument Realization', in Everaert, M., and van Riemsdijk, H. (eds.) *The Blackwell Companion to Syntax (Vol 4)*. Malden, MA: Blackwell, pp. 118-179.
- Dobrovie-Sorin, C. (2013) *A Reference Grammar of Romanian*. (Volume 207 of Linguistik Aktuell/Linguistics Today). John Benjamins Publishing Company.
- Doliana, A. (2013) 'The Super-Strong Person Case Constraint: Scarcity of Resources by Scale-Driven Impoverishment', in Heck, F., and Assmann, A. (eds.) *Rule Interaction in Grammar*. (Linguistische Arbeits Berichte 90). Universität Leipzig, pp. 177-202.
- Doron, E. (2003) 'Agency and Voice: The Semantics of the Semitic Templates', *Natural Language Semantics*, 11, pp. 1-67.
- Doron, E., and Rappaport Hovav, M. (2007) 'On the uniform application of valence changing operations', in Falk, Y. N. (ed.) *Proceedings of Israel Association for Theoretical Linguistics*. Available at: <http://linguistics.huji.ac.il/IATL/23/Doron-Rappaport.pdf> (Accessed: 4th August 2008).
- Doussinet, R. (1971) *Grammaire saintongeaise: Étude des structures d'un parler régional*. La Rochelle: Rupella.
- DRAE (2001) *Diccionario de la lengua española*. Available at: <http://www.rae.es/recursos/diccionarios/diccionarios-antiores-1726-2001/diccionario-de-la-lengua-espanola-2001> (Accessed: 28th June 2016).
- Dryer, M. S. (1997) 'Are Grammatical Relations Universal?', in Bybee, J., Haiman, J., and Thompson, S. A. (eds.) *Essays on Language Function and Language Type*. Amsterdam: Benjamins, pp. 115-143.
- Elman, J. L., Hare, M., and McRae, K. (2005) 'Cues, Constraints, and Competition in Sentence Processing', in Tomasello, M., and Slobin, D. I. (eds.) *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*. Mahwah, NJ: Lawrence Erlbaum, pp. 111-138.
- Embick, D. (2004a) 'Unaccusative Syntax and Verbal Alternations', in Alexiadou, A., Anagnostopoulou, E., and Everaert, M. (eds.) *The unaccusativity puzzle: explorations of the syntax-lexicon interface*. Oxford, UK: Oxford University Press, pp. 137-158.
- Embick, D. (2004b) 'On the structure of resultative participles in English', *Linguistic Inquiry*, 35, pp. 355-392.
- Enrique-Arias, A. (2011) 'El sincretismo de persona en el plural (os, nos y los)', *Lingüística Iberoamericana*, 43, pp. 99-118.
- Erteschik-Shir, N. (1979) 'Discourse constraints on dative movement', in Givón, T. (ed.) *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press.
- Erteschik-Shir, N. (1997) *The Dynamics of Focus Structure*. Cambridge.
- Espinal, M. T. (2004) 'Lexicalization of light verb structures and the semantics of nouns', *Catalan Journal of Linguistics*, 3, pp. 15-43.
- Espinal, M. T. (2009) 'Clitic incorporation and abstract semantic objects in idiomatic constructions', *Linguistics*, 47.6, pp. 1221-1271.
- Evans, J. K., Lepschy, G., Morris, S., Newman, J., and Waton, D. (1978) 'Italian clitic clusters', *Studi italiani di linguistica teorica e applicata*, 7, pp. 153-168.
- Everett, D. L. (2005) 'Cultural Constraints on Grammar and Cognition in Pirahã', *Current Anthropology*, 46, pp. 621-646.

- Fabra, P. (1956) *Gramàtica catalana. Prefaci i edició a cura de Joan Coromines*. Barcelona: Teide.
- Faltz, L. M. (1985) *Reflexivization: A Study in Universal Syntax*. New York: Garland.
- Farkas, D., and Kazasis, K. (1980) 'Clitic Pronouns and Topicality in Romanian', in Kreiman, J. and Ojeds, A. (eds.) *Papers from the 16th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago University Press, pp. 75-82.
- Feliu Arquiola, E. (2008) 'Spanish Middle Sentences: A Role and Reference Grammar Approach', in Kailuweit, R., Wiemer, B., Staudinger, E., and Matasović, R. (eds.) *New Applications of Role & Reference Grammar: Diachrony, Grammaticalization, Romance Languages*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 356-388.
- Fellbaum, C., and Zribi-Hertz, A. (1989) *The middle construction in French and English: A comparative study of its syntax and semantics*. Bloomington, Indiana: Indiana University Linguistics Club.
- Fernández Lagunilla, M. (1975) 'Acerca de la secuencia "se impersonal + enclítico de 3. a persona": ¿una restricción superficial?', *Revista Española de Lingüística*, 5.1, p. 177-193.
- Fernández Ramírez, S. (1964) 'Un proceso lingüístico en marcha', *Presente y futuro de la lengua española: actas de la Asamblea de Filología del I Congreso de Instituciones Hispánicas*. Madrid: Oficina Internacional de Información y Observación del Español, pp. 277-285.
- Fernández Ramírez, S. (1986) *Gramática española. El verbo y la oración*. Madrid: Arco Libros.
- Fernández Soriano, O. (1999) 'Two types of impersonal constructions in Spanish: locative and dative subjects', *Syntax*, 2.2, pp. 101-140.
- Fernández Soriano, O. (2000) 'Datives in construction with unaccusative se', *Catalan Working Papers in Linguistics*, 7, pp. 89-105.
- Fernández Soriano, O., and Mendikoetxea, A. (2011) 'Non-selected dative subjects in anticausative constructions', *Archivio glottologico italiano*, 96.1, pp. 87-127.
- Fernández-Ordóñez, I. (1993) 'Leísmo, láismo y loísmo: estado de la cuestión', in Fernández Soriano, O. (ed.) *Los pronombres átonos*. Madrid: Taurus, pp. 63-96.
- Fernández-Ordóñez, I. (1999) 'Leísmo, láismo y loísmo', in Bosque, I., and Demonte, V. (eds.) *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe, pp. 1317-97.
- Fernández-Ordóñez, I. (2009) 'The development of mass/count distinctions in Indo-European varieties', in Bubenik, V., Hewson, J., and Rose, S. (eds.) *Grammatical Change in Indo-European Languages. Papers Presented at the Workshop on Indo-European Linguistics at the XVIIIth International Conference on Historical Linguistics, Montreal 2007*. Amsterdam: John Benjamins, pp. 55-68.
- Ferrer, E. B. (1984) *Storia Linguistica della Sardegna*. Tübingen, Germany: Niemeyer.
- Fillmore, C. (1985) 'Frames and the semantics of understanding', *Quaderni di semantica*, 6.2.12, pp. 222-254.
- Finegan, E. (1995) 'Subjectivity and subjectification: An introduction', in Stein, D., and Wright, S. (eds.) *Subjectivity and Subjectivisation*. Cambridge: Cambridge University Press, pp. 1-15.
- Fisher, C. (2000) 'From form to meaning: a role for structural alignment in the acquisition of language', *Advances in Child Development and Behavior*, 27, pp. 1-53.

- Fleurent, C. (2015) *La phrase impérative en français formel et informel d'un point de vue syntaxique*. Masters Thesis. Université Laval, Québec, Canada.
- Flórez, L. (1977) *Apuntes de español*. Bogotá: Instituto Caro y Cuervo.
- Floricić, F., and Molinu, L. (2003) 'Imperativi "monosillabici" e "Minimal Word" in italiano "standard" e in sardo', in Giacomo-Marcellesi, M., and Rocchetti, A. (eds.) *Il verbo italiano: studi diacronici, sincronici, contrastive, didattici*. Roma: Bulzoni, pp. 343-357.
- Floricić, F., and Molinu, L. (2012) 'Romance monosyllabic imperatives and markedness', in Stolz, T., Nau, N., and Stroh, C. (eds.) *Monosyllables: from phonology to typology*. Berlin: Akademie Verlag, pp. 149-172.
- Fodor, J. (2001) 'Parameters and the periphery: reflections on Syntactic Nuts', *Journal of Linguistics*, 37, pp. 367-392.
- Fodor, J., and Lepore, E. (1998) 'The Emptiness of the Lexicon: Reflections on James Pustejovsky's The Generative Lexicon', *Linguistic Inquiry*, 29, pp. 269-288.
- Folli, R. (2000) 'Constructing Telicity in English and Italian', *Paper presented at the Workshop on Paths and Telicity in Event Structure. ESSLLI 2000, Birmingham*.
- Folli, R. (2003) *Constructing Telicity in English and Italian*. Doctoral Thesis. University of Oxford.
- Folli, R., and Harley, H. (2004) *Waltzing Matilda around and around: On the licensing of directed- motion resultatives*. Manuscript. University of Cambridge and University of Arizona.
- Folli, R., and Harley, H. (2005) 'Flavors of v: Consuming results in Italian and English', in Kempchinsky, P., and Slabakova, R. (eds.) *Aspectual Inquiries*, Dordrecht: Springer, pp. 95-120.
- Folli, R., and Harley, H. (2006) 'On the licensing of causatives of directed motion: Waltzing Matilda all over*', *Studia Linguistica*, 60.2, pp. 121-155.
- Fonagy, I. (1979) 'L'accent français: accent probabilitaire', *Studia Phonetica*, 15, pp. 123-133.
- Fontana, J. (1994) 'A variationist account of the development of the Spanish clitic system', in Beals, K. (ed.) *Papers from the 13th regional Meeting of the Chicago Linguistic Society, Vol. 2: The parasession on variation in linguistic theory*. Chicago: Chicago Linguistic Society, pp. 87-100.
- Fortescue, M. (2002) 'Report to the Elders (a discourse on semantic relativity)', *Acta Linguistica Hafniensia*, 34, pp. 107-163.
- Foulet, L. (1919) *Petite syntaxe de l'ancien français*. Paris: Champion.
- Foulet, L. (1921) 'Comment ont évolué les formes de l'interrogation', *Romania*, 47, pp. 243-348.
- Foulet, L. (1924) 'L'accent tonique et l'ordre des mots: formes faibles du pronom personnel après le verbe', *Romania*, 50, pp. 54-93.
- Francez, I. (2007) *Existential propositions*. Doctoral Thesis. Stanford University.
- Franco, J. (1993) *On Object Agreement in Spanish*. Doctoral Thesis. University of Southern California.
- Franco, J. (2000) 'Agreement as a Continuum, the Case of Spanish Pronominal Clitics', in Beukema, F. and den Dikken, M. (eds.) *Clitic Phenomena in European Languages*. Amsterdam: John Benjamins, pp. 147-190.

- Franco, J., and Huidobro, S. (2008) 'Ethical Datives, Clitic Doubling and the Theory of pro', in Bruhn de Garavito, J., and Valenzuela, E. (eds.) *Selected Proceedings of the 10th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 215-224.
- Freeze, R. (1992) 'Existentials and other locatives', *Language*, 68.3, pp. 553-595.
- Gabrielli, A. (1956) *Dizionario linguistico moderne*. Milano: Mondadori.
- Galambos, S. J. (1985) 'Mechanisms of change in the position of object pronouns: from classical Latin to Modern French', in King, L. D., and Maley, C. A. (eds.) *Selected papers from the XIIIth Linguistic symposium on Romance Languages (1983)*. John Benjamins, pp. 99-116.
- García Negroni, M. M. (2002) 'La construcción media con se', in Sánchez López, C. (ed.) *Las construcciones con se*. Madrid: Visor, pp. 275-308.
- García, E. C. (1975) *The role of theory in linguistic analysis: the Spanish pronoun system*. New York: American Elsevier Pub. Co.
- García, E. C. (2009) *The motivated syntax of arbitrary signs cognitive constraints on Spanish clitic clustering*. Amsterdam/Philadelphia: John Benjamins Pub. Co. Available at: <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622874> (Accessed: 22nd October 2012).
- García, E. C., and van Putte, F. (1987) 'Forms are silver, nothing is gold', *Folia Linguistica Historica VIII*, pp. 365-384.
- Garrod, S., and Pickering, M. J. (1999) 'Issues in language processing', in Garrod, S., and Pickering, M. J. (eds.) *Language processing*. Hove: Psychology Press, pp. 1-11.
- Gast, V., and Haas, F. (2008) 'On Reciprocal and Reflexive Uses of Anaphors in German and Other European Languages', in König, E., and Gast, V. (eds.) *Reciprocals and Reflexives: Cross-Linguistic and Theoretical Explorations*. Berlin: Mouton de Gruyter, pp. 307-346.
- Gavarró, A. (1992) 'Empty Objects in Catalan', *Catalan Working Papers in Linguistics 1992*, pp.145-161.
- Geniušienė, E. (1987) *The Typology of Reflexives*. (Empirical Approaches to Language Typology 2). Berlin: Mouton de Gruyter.
- Gentner, D. (1983) 'Structure-Mapping: a Theoretical Framework for Analogy', *Cognitive Science*, 7, pp. 155-170.
- Gentner, D. (1988) 'Metaphor as Structure Mapping: The Relational Shift', *Child Development*, 59, pp. 47-59.
- Gentner, D. (2003) 'Why We're So Smart', in Gentner, D., and Goldin-Meadow, S. (eds.) *Language in Mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press, pp. 195-235.
- Gentner, D., and Goldin-Meadow, S. (eds.) (2003) *Language in Mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.
- Gentner, D., and Markman, A. B. (1997) 'Structure mapping in Analogy and Similarity', *American Psychologist*, 52, pp. 45-56.
- Gentner, D., Holyoak, K. J., and Kokinov, B. N. (eds.) (2001) *The Analogical Mind*. Cambridge, MA: MIT Press.

- Georgi, D. (2008) 'A Distributed Morphology Approach to Argument Encoding in Kambera', *Linguistische Berichte*, 213, pp. 45-63.
- Gerlach, B. (2002) *Clitics between syntax and lexicon*. Amsterdam/Philadelphia: J. Benjamins. Available at: <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622308> (Accessed: 19th October 2012).
- Giacomo-Marcellesi, M. (1997) *Corse*. München/Newcastle: LINCOM Europa.
- Gili Gaya, S. (1943) *Curso superior de sintaxis española*. Barcelona: Vox.
- Gili Gaya, S. (1964) *Curso superior de sintaxis española*. Barcelona: Biblograf.
- Giurgea, I. (2012) 'The "pro-N" clitic en / ne and the theory of cliticization', *Bucharest Working Papers in Linguistics*, 1, pp. 39-56.
- Giurgea, I. (2013) 'L'origine Des Clitiques Roumains de 3^e Personne Pluriel Datif et de 1^{ère} Personne Pluriel Datif-Accusatif', *RRL*, LVIII.2, pp. 125-136.
- Givón, T. (ed.) (1979) *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press.
- Givón, T. (1984) *Syntax: a functional typological-introduction*. Amsterdam: John Benjamins.
- Givón, T. (1990) *Syntax – A Functional-Typological Introduction, volume II*. Amsterdam: John Benjamins.
- Givón, T., and Yang, L. (1994) 'The rise of the English GET-passive', in Fox, B., and Hopper, P. (eds.) *Voice, form and function*. (Typological Studies in Language 27). Amsterdam: John Benjamins, pp. 119-150.
- Goldberg, A. E. (1995) *Constructions : a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2002) 'Surface generalizations: An alternative to alternations', *Cognitive Linguistics*, 13, pp. 327-356.
- Gómez Torrego, L. (1992) *Valores gramaticales de se*. Madrid: Arco Libros.
- Goosse, A. (2000) 'Evolution de la syntax', in Antoine, G., and Cerquiglini, B. (eds.) *Histoire de la langue française 1945–2000*. Paris: CNRS, pp. 107-145.
- Goria, C. (2004) *Subject Clitics in the Northern Italian Dialects: A Comparative Study Based on the Minimalist Program and Optimality Theory*. Dordrecht: Kluwer.
- Grashchenkov, P. and Markman, V. G. (2008) 'Non-Core Arguments in Verbal and Nominal Predication: High and Low Applicatives and Possessor Raising', in Abner, N., and Bishop, J. (eds.) *Proceedings of the 27th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 185-193.
- Green, D. M. (2004) 'Review of Postverbal behavior, by Thomas Wasow', *Language*, 80, pp. 327-331.
- Green, L. J. (2002) *African American English: a linguistic introduction*. Cambridge; New York: Cambridge University Press.
- Grevisse, M., and Goosse, A. (2008) *Le bon usage: grammaire française*. Bruxelles [etc.]: De Boeck.

- Gries, S. T. (2003) *Multifactorial analysis in corpus linguistics: a study of particle placement*. London: Continuum.
- Gries, S. T., and Stefanowitsch, A. (2004) 'Extending collocation analysis. A corpus-based perspective on 'alternations'', *International Journal of Corpus Linguistics*, 9, pp. 97-129.
- Grimalt, P. (2002) *Enclisi i accentuació en el català de Mallorca i Menorca*. Doctoral Thesis. Universitat de Barcelona.
- Grimshaw, J. (1982) 'On the lexical representation of Romance Reflexive Clitics', in Bresnan, J. (ed.) *The mental representation of grammatical relations*. Cambridge, MA: MIT Press, pp. 87-148.
- Grimshaw, J. (1990) *Argument Structure*. Cambridge, Mass: MIT Press.
- Grimshaw, J. (1997) 'The Best Clitic: Constraint Conflict in Morphosyntax', in Haegeman, L. (ed.) *Elements of Grammar*. Dordrecht: Kluwer, pp. 169-196.
- Grimshaw, J. (2001) 'Optimal Clitic Positions and the Lexicon in Romance Clitic Systems', in Legendre, G., Grimshaw, J., and Vikner, S. (eds.) *OT Syntax*. MIT Press, pp. 205-240.
- Gross, M. (1968) *Grammaire transformationnelle du français: syntaxe du verbe*. Paris: Larousse.
- Gutiérrez Ordóñez, S. (1999) 'Los dativos', in Bosque, I., and Demonte, V. (eds.) *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe, pp. 1855-1930.
- Gutiérrez Ordóñez, S. (2002) *Forma y sentido en sintaxis*. Madrid: Arco Libros.
- Hagedorn, C. (2009) *True Clitics and Weak Pronouns in Seneghese*. Stony Brook University.
- Haiman, J. (1980) 'The iconicity of grammar: Isomorphism and motivation', *Language*, 54.3, pp. 515-540.
- Haiman, J. (1983) 'Iconic and Economic Motivations', *Language*, 59, pp. 781-819.
- Hale, K., and Keyser, S. J. (1986) *Some Transitivity Alternations in English*. (=Lexicon Project Working Papers 7). Cambridge, MA: Center for Cognitive Science, MIT.
- Hale, K., and Keyser, S. J. (1987) *A view from the Middle*. (=Lexicon Project Working Papers 10). Cambridge MA: Center for Cognitive Science, MIT.
- Hall, R. (1960) 'Statistica grammaticale. L'uso di gli e loro come regime indiretto', *Lingua nostra*, 21, pp. 58-65.
- Hall, R. (1965) 'The "neuter" in Romance: a pseudo problem', *Word*, 21.3, pp. 421-427.
- Halle, M. (2000) 'Distributed morphology: Impoverishment and fission', in Lecarme, J., Lowenstamm, J., and Shlonsky, U. (eds) *Research in Afroasiatic Grammar: Papers from the Third conference on Afroasiatic Languages*. Amsterdam: Benjamins, pp. 125-150.
- Harley, H. (1995) *Subjects, events and licensing*. Doctoral Thesis. MIT.
- Harley, H. (1998) 'You're having me on! Aspects of have', in Gueron, J., and Zribi-Hertz, A. (eds.) *La grammaire de la possession*. Nanterre: Publidix, pp. 195-226.
- Harley, H. (2002) 'Possession and the double object construction', in Pica, P., and Rooryck, J. (eds.) *Yearbook of Linguistic Variation, vol 2*. Amsterdam: John Benjamins, pp. 31-70.

- Harley, H. (2004) 'Wanting, having and getting. A note on Fodor and Lepore 1998', *Linguistic Inquiry*, 35, pp. 255-267.
- Harris, J. (1996) 'The morphology of Spanish clitics', in Campos, H., and Kempchinsky, P. (eds.) *Evolution and Revolution in Linguistic Theory: Essays in Honour of Carlos Otero*. Washington D. C.: Georgetown University Press, pp. 168-197.
- Harris, J. (1997) 'Why N'HO is pronounced [LI] in Barceloni Catalan: Morphological impoverishment, merger, fusion, and fission', *Recherces linguistiques de Vincennes*, 26, pp. 61-86.
- Harris, J., and Halle, M. (2005) 'Unexpected Plural Inflections in Spanish: Reduplication and Metathesis', *Linguistic Inquiry*, 36.2, pp. 195-222.
- Härtl, H. (2003) 'Conceptual and grammatical characteristics of argument alternations: The case of decausative verbs', *Linguistics*, 41, pp. 883-916.
- Hasher, L., and Zacks, R. T. (1984) 'Automatic Processing of Fundamental Information. The Case of Frequency of Occurrence', *American Psychologist*, 39, pp. 1372-1388.
- Haspelmath, M. (1993) 'More on the typology of inchoative/causative verb alternations', in Comrie, B., and Polinsky, M. (ed.) *Causatives and Transitivity*. Amsterdam: John Benjamins, pp. 87-120.
- Haspelmath, M. (2000) 'Why can't we talk? A review article of Newmeyer (1998)', *Lingua*, 110.4, pp. 235-255.
- Haspelmath, M. (2004) 'Explaining the Ditransitive Person-Role Constraint: A Usage-Based Account', *Constructions 2*. Available at: <http://www.constructions-online.de> (Accessed: 1st January 2016).
- Hatcher, A. G. (1942) *Reflexive Verbs. Latin, Old French, Modern French*. (The John Hopkins Studies in Romance Literatures and Languages 43). Baltimore: The John Hopkins Press.
- Hawkins, J. A. (2004) *Efficiency and complexity in grammars*. Oxford/New York: Oxford University Press. Available at: <http://site.ebrary.com/id/10263703> (Accessed: 21st October 2012).
- Hay, J., Kennedy, C., and Levin, B. (1999) 'Scalar structure underlies telicity in degree achievements', in Matthews, T., and Strolovich, D. (eds.) *Proceedings of SALT 9*. Ithaca, NY: CLC Publications, pp. 127-144.
- Hazout, I. (2004) 'The syntax of existential constructions', *Linguistic Inquiry*, 35, pp. 393-430.
- Heap, D. (2003) 'Secuencias « invertidas » de clíticos: un cambio (¿?) En tiempo real', in Girón Alconchel, J. L., and Bustos Tovar, J. J. (eds.) *Actas del VI Congreso Internacional de Historia de la Lengua española : Madrid, 29 de septiembre-3 octubre 2003*. Available at: <https://dialnet.unirioja.es/servlet/libro?codigo=8863> (Accessed: 18th July 2016).
- Heap, D., and Roberge, Y. (2001) 'Cliticisation et théorie syntaxique, 1971-2001', *Revue Québécoise De Linguistique*, 30.1, pp. 63-90.
- Heggie, L. A., and Ordóñez, F. (eds.) (2005) *Clitic and affix combinations : theoretical perspectives*. Amsterdam [etc.]: John Benjamins.
- Heidinger, S., and Schäfer, F. (2008) 'On the French Reflexive Passive and Anticausative. A diachronic view from the *par*-phrase', in Fagard, B., Prevost, S., Combettes, B., and Bertrand, O. (eds.) *Etudes de linguistique diachronique*. Bern: Peter Lang, pp. 135-152.
- Heine, B. (1992) 'Grammaticalization chains', *Studies in Language*, 16.2, pp. 335-368.

- Heine, B. (1997) *Cognitive foundations of grammar*. New York: Oxford University Press.
- Heine, B. (2000) 'Grammaticalization chains across languages: An example from Khoisa', in Gildea, S. (ed.) *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*. (Typological Studies in Language 43). Amsterdam/Philadelphia: John Benjamins, pp. 177-199.
- Heine, B., and Miyashita, H. (2008) 'The Intersection Between Reflexives and Reciprocals: A Grammaticalization Perspective', in König, E., and Volker, G. (eds.) *Reciprocals and Reflexives: Cross-Linguistic and Theoretical Explorations*. Berlin: Mouton de Gruyter, pp. 169-224.
- Heine, B., Ulrike, C., and Hünemeyer, F. (1991) *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hernández Alonso, C. (1966) 'Del se reflexivo al impersonal', *Archivum XVI*, pp. 39-66.
- Herschensohn, J. (1992) 'Case Marking and French Psych-Verbs', *LI Lingvisticae Investigationes*, 16.1, pp. 21-40.
- Hetzron, R. (1977) 'Clitic pronouns and their linear representation', *Forum Linguisticum*, 1, pp. 189-215.
- Higginbotham, J. (1997) *Location and causation*. Masters Thesis. University of Oxford.
- Himmelmann, N. P. (2004) 'Lexicalization and grammaticalization: Opposite or orthogonal?', in Bisang, W., Himmelmann, N. P., and Wiemer, B. (eds.) *What Makes Grammaticalization: A Look at Its Fringes and Its Components*. (Trends in Linguistics. Studies and Monographs 158). Berlin/New York: Mouton de Gruyter, pp. 21-41.
- Hoekstra, T. (1995) 'To have to be dative', in Haider, H., Olsen, S., and Vikner, S. (eds.) *Studies in comparative Germanic syntax*. Dordrecht: Kluwer Academic Publishers, pp. 119-137.
- Hollmann, W. (2005) 'The iconicity of infinitival complementation in Present-day English causatives', in Maeder, C., Fischer, O., and Herlofsky, W. (eds.) *Outside-in - Inside-out. Iconicity in Language and Literature*, 4. Amsterdam: John Benjamins, pp. 287-306.
- Holtus, G., Metzeltin, M., and Schmitt, C. (1989) *Lexikon der Romanistischen Linguistik III*. Tübingen: Max Niemeyer.
- Hopper, P., and Traugott, E. C. (2003) *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.
- Horn, L. R. (1985) 'Metalinguistic negation and pragmatic ambiguity', *Language*, 61, pp. 121-174.
- Horn, L. R. (2008) "'I love me some him": The landscape of non-argument datives', in Bonami, O., and Cabredo Hofherr, P. (eds.) *Empirical Issues in Syntax and Semantics*, 7, pp. 169-192.
- Hornstein, N. (2001) *Move! A Minimalist Theory of Construal*. Blackwell Publishers.
- Horvath, J., and Siloni, T. (2011) 'Anticausatives: Against reflexivization', *Lingua*, 121, pp. 2176-2186.
- Horvath, J., and Siloni, T. (2013) 'Anticausatives have no cause(r): A rejoinder to Beavers and Koontz-Garboden', *Lingua*, 131, pp. 217-230.
- Hourcade, A. (1986) *Grammaire béarnaise*. Pau: Los Caminaires.

- Howell, E. G., and Howell, J. (eds.) (2006) *Proceedings of Semantic and Linguistic Theory XV*. Ithaca, NY: Cornell University Press.
- Hualde, J. I. (2003) 'El modelo métrico y autosegmental', in Prieto, P. (ed.) *Teorías de la entonación*. Barcelona: Editorial Ariel, pp. 155-184.
- Huang, C. T. J. (1999) 'Chinese passives in comparative perspective', *The Tsing Hua Journal of Chinese Studies*, 29.4, pp. 423-509.
- Hudson, R., Rosta, A., Holmes, J., and Gisborne, N. (1996) 'Synonyms and syntax', *Journal of Linguistics*, 32, pp. 439-446.
- Hulk, A. (1989) 'La construction impersonnelle et la structure de la phrase', *Recherches Linguistiques de Vincennes*, 18, pp. 59-81.
- Hurford, J. R. (2003) 'The Language Mosaic and its Evolution', in Christensen, K., and Kirby, S. (eds.) *Language evolution*. Oxford: Oxford University Press, pp. 38-57.
- Hurst, D. A. (1951) 'Influence of Subject and Connotation of Force', *Hispania*, 34, pp. 74-8.
- Jackendoff, R. (1992) 'Mme Tussaud meets the binding theory', *Natural Language and Linguistic Theory*, 10, pp. 1-31.
- Jackendoff, R. (1997) *The architecture of the language faculty*. (Linguistic inquiry monographs, 28). Cambridge, Mass: MIT Press.
- Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: OUP.
- Jaeggli, O. (1982) *Topics in Romance Syntax*. Dordrecht: Foris.
- Jensen, F. (1986) *The Syntax of Medieval Occitan*. (Beihefte Zur Zeitschrift für Romanische Philologie, 208). Tübingen: Max Niemeyer Verlag.
- Jespersen, O. (1924) *The Philosophy of Grammar*. London: George Allen & Unwin Ltd.
- Jones, M. A. (1993) *Sardinian syntax*. London/New York: Routledge.
- Jones, M. A. (1996) *Foundations of French syntax*. Cambridge: Cambridge University Press.
- José, B., and Auger, J. (2005) 'Geminates and Picard Pronominal Clitic Allomorphy', *Catalan Journal of Linguistics*, 4, pp. 127-154.
- Jouitteau, M., and Řezáč, M. (2007) 'The French ethical dative: 13 syntactic tests', *Bucharest Working Papers in Linguistics*, IX, pp. 97-108.
- Jun, S.-A., and Fougeron, C. (2000) 'A phonological model of French intonation', in Botinis, A. (ed.) *Intonation: Analysis, modelling and technology*. Boston: Kluwer, pp. 209-242.
- Jurafsky, D. (2003) 'Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production', in Bod, R., Hay, J., and Jannedy, S. (eds.) *Probabilistic Linguistics*. Cambridge, MA: MIT Press, pp. 39-95.
- Kager, R. (1995) 'Consequences of catalexis', in van der Hulst, H. and van de Weijer, J. (eds.) *Leiden in Last. HIL. Phonology Papers I*. The Hague: Holland Academic Graphics, pp. 269-298.

- Kaisse, E. (1985) *Connected Speech. The Interaction of Syntax and Phonology*. San Diego, CA: Academic Press.
- Kako, E. (2006) 'Thematic role properties of subjects and objects', *Cognition*, 101, pp. 1-42.
- Kallulli, D. (1999) 'Non-active morphology in Albanian and event (de)composition', in Kenesei, I. (ed.) *Crossing Boundaries*. (Current Issues in Linguistic Theory 182). Amsterdam: John Benjamins, pp. 263-292.
- Kallulli, D. (2007) 'Rethinking the passive/anticausative distinction', *Linguistic Inquiry*, 38, pp. 770-778.
- Kamp, H., and Roßdeutscher, A. (1994) 'Remarks on Lexical Structure and DRS Construction', *Theoretical Linguistics*, 20, pp. 97-164.
- Kany, C. (1951) *American-Spanish Syntax*. Chicago: University of Chicago Press.
- Kany, C. (1969) *Sintaxis hispanoamericana*. Madrid: Gredos.
- Kay, P. (1997) *Words and the Grammar of Context*. Stanford, CA: CSLI Publications.
- Kayne, R. (1975) *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.
- Kayne, R. S. (1977) *Syntaxe du français: le cycle transformationnel*. Paris: Éditions du Seuil.
- Kayne, R. S. (1994) *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press. Available at: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk &AN=48755> (Accessed: 6th November 2014).
- Kayne, R. S. (2000) *Parameters and universals*. Oxford; New York: Oxford University Press.
- Kayne, R. S. (2008) 'Expletives, Datives, and the Tension between Morphology and Syntax', in Biberauer, T. (ed.) *The Limits of Syntactic Variation*. Amsterdam/Philadelphia, PA: John Benjamins, pp. 175-217.
- Keller, R. (1995) *Zeichentheorie*. Tübingen: Francke.
- Kemmer, S. (1993) *The middle voice*. Amsterdam; Philadelphia: John Benjamins Pub. Co. Available at: <http://site.ebrary.com/id/10509465> (Accessed: 6th November 2014).
- Kemmer, S. (1994) 'Middle voice, transitivity, and the elaboration of events', in Fox, B., and Hopper, P. J. (eds.) *Voice: Form and function*. (Typological Studies in Language 27). Amsterdam/Philadelphia: John Benjamins, pp. 179-230.
- Kim, M., and Repetti, L. (2013) 'Bitonal pitch accent and phonological alignment in Sardinian', *Probus*, 25.2, pp. 267-300.
- Kinder, J. J., and Savini, V. M. (2004) *Using Italian: a guide to contemporary usage*. Cambridge/New York: Cambridge University Press.
- Kiparsky, P. (1971) 'Historical Linguistics', in Dingwall, W. O. (ed.) *A survey of linguistic science*. College Park, MD: University of Maryland, pp. 576-649.
- Kirby, S. (1998) 'Fitness and the selective adaptation of language.', in Hurford, J. R., Studdert-Kennedy, M., and Knight, C. (eds.) *Approaches to the evolution of language*. Cambridge: Cambridge University Press, pp. 359-383.

- Kirby, S., and Brighton, H. (2004) 'From UG to Universals: linguistic adaptation through iterated learning', *Studies in Language*, 28, pp. 587-607.
- Klausenburger, J. (2000) *Grammaticalization: Studies in Latin and Romance Morphosyntax*. (Current Issues in Linguistic Theory 193). Amsterdam/Philadelphia: John Benjamins.
- Klavans, J. L. (1982) *Some Problems in Theory of Clitics*. Indiana Univ.: Linguistics Club.
- Klavans, J. L. (1985) 'The Independence of Syntax and Phonology in Cliticization', *Language*, 61.1, pp. 95-120.
- Kleiber, G. (1993) 'Iconicité d'isomorphisme et grammaire cognitive', *Faits de langue*, 1, pp. 105-21.
- Kliffner, M. D. (1983) 'Beyond Syntax: Spanish Inalienable Possession', *Linguistics*, 21, pp. 759-94.
- Koch, P. (2004) 'Metonymy between pragmatics, reference, and diachrony', *Metaphorik.de*. Available at: http://www.metaphorik.de/sites/www.metaphorik.de/files/journal-pdf/07_2004_koch.pdf (Accessed: 1st January 2016).
- Kok, D. A. (1985) *La place du pronom personnel régime conjoint en français : un étude diachronique*. Amsterdam: Rodopi, B.V.
- Koontz-Garboden, A. (2007) *States, changes of state, and the Monotonicity Hypothesis*. Doctoral thesis. Stanford University.
- Koontz-Garboden, A. (2009) 'Anticausativization', *Natural Language and Linguistic Theory*, 27, pp. 77-138.
- Kratzer, A. (1995) 'Stage Level and Individual Level Predicates', in Carlson, G., and Pelletier, F. J. (eds.) *The Generic Book*. Chicago, IL: The University of Chicago Press, pp. 125-175.
- Kuhn, A. (2008) *El dialecto altoaragonés*. Zaragoza: Xordiga.
- Kuikenheim, L. (1968) *La grammaire historique de la langue français. Les Syntagmes*. Leiden: Universitaire Pers Leiden.
- La Fauci, N., and Loporcaro, M. (1993) 'Grammatical relations and syntactic levels in Bonorvese morphosyntax', in Belletti, A. (ed.) *Syntactic Theory and the Dialects of Italy*. Turin: Rosenberg & Sellier, pp. 155-203.
- La Fauci, N., and Loporcaro, M. (1997) 'Outline of a theory of existentials on evidence from Romance', *Studi Italiani di Linguistica Teorica e Applicata*, 26, pp. 5-55.
- Labelle, M. (1990) 'Unaccusatives and pseudo-unaccusatives in French', in Carter, J., Dechaine, R.-M., Philip, B. and Sherer, T. (eds.) *Proceedings of NELS 20*. Amherst: GLSA, pp. 303-317.
- Labelle, M. (1992) 'Change of state and valency', *Journal of Linguistics*, 28, pp. 375-414.
- Labelle, M. (2008) 'The French Reflexive and Reciprocal se', *Natural Language & Linguistic Theory*, 26.4, pp. 833-876.
- Labov, W. (1994) *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.

- Laca, B. (2001) 'Gramaticalización y variabilidad: propiedades inherentes y factores contextuales en la evolución del acusativo preposicional en español', in Wesch, A., Weidenbusch, W., Kailuweit, R., and Laca, B. (eds.) *Sprachgeschichte als Varietätengeschichte. Festschrift Jens Lüdtko*. Tübingen: Stauffenburg, pp. 195-203.
- Ladd, D. R. (1996) *Intonational phonology*. Cambridge: Cambridge University Press.
- Laenzlinger, C. (1993) 'A syntactic view of Romance pronominal sequences', *Probus*, 5.3, pp. 241-270.
- Laenzlinger, C. (1994) 'Enclitic clustering: The case of French positive imperatives', *Rivista di Grammatica Generativa*, 19, pp. 71-104.
- Lagae, V. (2002) 'Le passif pronominal: une forme complémentaire du passif périphrastique?', in Lagae, V., Carlier, A., and Benninger, C. (eds.) *Cahiers Chronos 10, Temps et aspect: de la grammaire au lexique*. Amsterdam: Rodopi, pp. 133-149.
- Laka, I. (1990) *Negation in syntax--on the nature of functional categories and projections*. Massachusetts: MIT.
- Lakoff, G. (1987) *Women, Fire, and Dangerous Things*. Chicago: Chicago University Press.
- Lambrecht, K. (1994) *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
- Lamiroy, B. (2003) 'Grammaticalization and external possessor structures in Romance and Germanic languages', in Coene, M., and D'Hulst, Y. (eds.) *From NP to DP. Volume II: the expression of possession in noun phrases*. Amsterdam: Benjamins, pp. 257-80.
- Landa Buil, M. (2005) 'Las Combinaciones De Clíticos En El Cheso', *Alazet*, 17, pp. 113-133.
- Langacker, R. W. (1987) *Foundations of cognitive grammar. Vol.1: Theoretical prerequisites*. Stanford, Calif: Stanford University Press.
- Langacker, R. W. (1990) 'Subjectification', *Cognitive Linguistics*, 1.1, pp. 5-38.
- Langacker, R. W. (1993) 'Reference-point constructions', *Cognitive Linguistics*, 4, pp. 1-38.
- Langacker, R. W. (2005) 'Construction Grammars: Cognitive, Radical, and Less So', in Ruiz de Mendoza Ibáñez, F., and Peña Cervel, M. S. (eds.) *Cognitive linguistics: Internal Dynamics and Interdisciplinary Interaction*. Berlin: Mouton de Gruyter, pp. 101-159.
- Lapesa, R. (1981) *Historia de la lengua española*. Madrid: Gredos.
- Lasnik, H. (1988) 'Subjects and the th-Criterion', *Natural Language and Linguistic Theory*, 6, pp. 1-17.
- Lázaro Carreter, F. (1964) 'Problemas de terminología lingüística', *Presente y futuro de la lengua española, vol. II*. Madrid: Ediciones Cultura Hispánica, pp. 383-392.
- Leclère, C. (1976) 'Datifs syntaxiques et datifs éthiques', in Chevalier, J. C., and Gross, M. (eds.) *Méthodes en grammaire française*. Paris: Klincksieck, pp. 73-96.
- Ledgeway, A. (2008) 'Sulla storia dei verbi copulari dei dialetti dell'alto Meridione: Il caso del napoletano', *The Italianist*, 28, pp. 281-301.
- Ledgeway, A. (2009) *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.

- Leech, G., and Svartvik, J. (1994) *A communicative grammar of English*. London: Longman.
- Lehmann, C. (1995) *Thoughts on Grammaticalization*. München/Newcastle: Lincom Europa.
- Lehmann, C. (2002) 'New reflections on grammaticalization and lexicalization', in Wischer, I., and Diewald, G. (eds.) *New Reflections on Grammaticalization*. (Typological Studies in Language 49). Amsterdam/Philadelphia: John Benjamins, pp. 1-18.
- Lehmann, C. (2004) 'Theory and method in grammaticalization', *Zeitschrift für German- istische Linguistik*, 32.2, pp. 152-187.
- Lekakou, M. (2005) *In the middle, somewhat elevated*. Doctoral Thesis. UCL.
- Lenz, R. (1935) *La oración y sus partes*. Madrid: Centro de Estudios Históricos.
- Leonetti, M. (2004) 'Specificity and differential object marking in Spanish', *Catalan Journal of Linguistics*, 3, pp. 75-114.
- Leonetti, M. (2005) 'Definiteness effects and the role of the coda in existential constructions'. Available at: <http://www2.uah.es/leonetti/papers/Def&Coda.pdf> (Accessed: 1st January 2016).
- Leonetti, M. (2008) 'Specificity in clitic doubling and in DOM', *Probus*, 20, pp. 33-65.
- Lepschy, A. L., and Lepschy, G. (1984) *La lingua italiana: storia, varietà dell'uso, grammatica*. Milano: Bompiani.
- Lepschy, G. (1989) 'Teoria e descrizione nella sintassi italiana. Reggenza, legamento e un uso di ne', *Romance Philology*, 42, pp. 422-35.
- Lespy, V. (1880) *Grammaire béarnaise*. 2nd edition. Paris/Maisonneuve/Geneva: Slatkine Reprints.
- Levin, B., and Rappaport Hovav, M. (1994) 'A Preliminary Analysis of Causative Verbs in English', *Lingua*, 92, pp. 35-77.
- Levin, B., and Rappaport Hovav, M. (1995) *Unaccusativity: at the syntax-lexical semantics interface*. Cambridge, Mass.: MIT Press.
- Levinson, S. C. (2000) *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Levinson, S. C. (2001) 'Covariation between spatial language and cognition, and its implications for language learning', in Bowerman, M., and Levinson, S. C. (eds.) *Language acquisition and conceptual development*. Cambridge: Cambridge University Press, pp. 566-588.
- Lewis, M. (1989) 'Aspect Marking in Gēgbe Prepositions: A Cognitive Approach to Multicategoriality', *Papers from the regional Meeting of the Chicago Linguistics Society*, 25, pp. 273-287.
- Lidz, J. (1997) 'When is a reflexive not a reflexive? Near-reflexivity and Condition R', in Kusumoto, K. (ed.) *NELS 27: Proceedings of the 27 th annual meeting of the North East Linguistics Society*. Amherst: GLSA, pp. 251-261.
- Lidz, J. (2001) 'Condition R', *Linguistic Inquiry*, 32.1, pp. 123-140.
- Lindblom, B. (1999) 'Emergent phonology', *Berkeley Linguistics Society*, 25, pp. 195-209.

- Lindblom, B., Guion, S., Hura, S., Moon, S., and Willerman, R. (1995) 'Is sound change adaptive?', *Rivista di Linguistica*, 7, pp. 5-37.
- Linder, K. P. (1987) *Grammatische Untersuchungen zur Charakteristik des Rätoromanschen in Graubünden*. Tübingen: Gunter Narr.
- Lohse, B., Hawkins, J. A., and Wasow, T. (2004) 'Domain minimization in English verb-particle constructions', *Language*, 80, pp. 238-261.
- Longobardi, G. (1991) 'Extraction from NP and the proper notion of head government', in Giorgi, A., and Longobardi, G. (eds.) *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press, pp. 57-112.
- Longobardi, G. (1994) 'Reference and proper names: A theory of N-movement in syntax and Logical Form', *Linguistic Inquiry*, 25, pp. 609-665.
- Lonzi, L. (1986) 'Pertinenza della struttura tema-rema per l'analisi sintattica', in Stammerjohann, H. (ed.) *Theme-Rheme in Italian*. Tübingen: Narr., pp. 99-120.
- López Del Castillo, L. (1976) *Llengua standard i nivells de llenguatge*. Barcelona: Laia.
- Loporcaro, M. (1995) 'Un capitolo di morfologia storica italo-romanza: it. ant. ne "ci" e forme meridionali congeneri', *L'Italia dialettale*, 58, pp. 1-48.
- Loporcaro, M. (1998) *Sintassi comperata dell'accordo participiale romanzo*. Torino: Rosenberg & Sellier.
- Loporcaro, M. (2000) 'Stress Stability Under Cliticization and the Prosodic Status of Romance Clitics', in Repetti, L. (ed.) *Phonological Theory and the Dialects of Italy*. Amsterdam/Philadelphia: John Benjamins, pp. 137-168.
- Loucks, J. T., and Baldwin, D. (2006) 'When is a Grasp a Grasp? Characterizing Some Basic Components of Human Action Processing', in Hirsh-Pasek, K., and Golinkoff, R.M. (eds.) *Action meets words: How children learn verbs*. New York: Oxford University Press, pp. 228-261.
- Lucy, J. A. (1992) *Grammatical categories and cognition : a case study of the linguistic relativity hypothesis*. Cambridge [etc.]: Cambridge University Press.
- Lucy, J. A., and Gaskins, S. (2003) 'Interaction of Language Type and Referent Type in the Development of Nonverbal Classification Preferences', in Gentner, D., and Goldin-Meadow, S. (eds.) *Language in Mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press, pp. 465-492.
- Lüdke, H. (1979) *Lucania*. Pisa: Pacini.
- Lüdtke, H. (ed.) (1980) *Kommunikationstheoretische Grundlagen des Sprachwandels*. Berlin: Mouton de Gruyter.
- Lyons, J. (1968) 'Existential, locative and possessive constructions', in Lyons, J. (ed.) *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press, pp. 495-504.
- MacDonald, J. E. (2004) 'Spanish Reflexive Pronouns: a Null Preposition Hypothesis', in Garding, G., and Tsujimura, M. (eds.) *WCCFL 23 Proceedings*. Somerville, MA: Cascadia Press, pp. 101-114.
- MacDonald, J. E., and Huidobro, S. (2010) 'The lack of Spanish Non-Argumental Clitic Doubling', in Borgonovo, C., Español-Echevarría, M., and Prévost, P. (eds.) *Selected Proceedings of the 12th Hispanic Linguistics Symposium*. Somerville, MA: Cascadia Proceedings Project, pp. 50-62.

- MacDonald, M. C., Perlmutter, N. J., and Seidenberg, M. S. (1994) 'Lexical Nature of Syntactic Ambiguity Resolution', *Psychological Review*, 101, pp. 676-703.
- Mackenzie, I. E. (2006) *Unaccusative verbs in Romance languages*. Basingstoke/New York: Palgrave Macmillan.
- MacLaury, R. E. (1991) 'Social and cognitive motivations of change: measuring variability in color semantics', *Language*, 67, pp. 34-62.
- Maiden, M. (1995) *A linguistic history of Italian*. London/New York: Longman.
- Maiden, M. (2004) 'When lexemes become allomorphs--On the genesis of suppletion', *Folia Linguistica*, XXXVIII, pp. 227-256.
- Maiden, M., and Robustelli, C. (2000) *A Reference Grammar of Modern Italian*. London: Edward Arnold.
- Maldonado, R. (1988) 'Energetic Reflexives in Spanish', in Axmaker, S., Jaisser, A., and Singmaster, H. (eds.) *General Session and Parasession on Grammaticalization. Proceedings of the Fourteenth Annual Meeting of Berkeley Linguistics Society, February 13-15, 1998*. Berkeley: Berkeley Linguistics Society, pp. 35-55.
- Maldonado, R. (1992) *Middle Voice: the case of Spanish 'se'*. Doctoral Thesis. University of California.
- Maldonado, R. (1993) 'Dynamic construals in Spanish', *Studi italiani di linguistica teorica e applicata*, XXII-3, pp. 531-566.
- Maldonado, R. (1999) *A media voz. Problemas conceptuales del clítico se*. México D.F: Universidad Nacional Autónoma de México.
- Maldonado, R. (2000) 'Conceptual Distance and Transitivity Increase in Spanish Reflexives', in Frajzyngier, Z. (ed.) *Reflexives Form and Function*. Amsterdam: Benjamins, pp. 153-186.
- Maldonado, R. (2002a) 'Objective and subjective datives', *Cognitive Linguistics*, 13, pp. 1-65.
- Maldonado, R. (2002b) 'Causatividad e (in)dependencia en el español de México', *Anuario de Letras*, XXXVIII. México: Instituto de investigaciones Filológicas, UNAM, pp. 281-308.
- Maldonado, R. (2008) 'Spanish Middle Syntax. A Usage Based Proposal for Grammar Teaching', in de Knop, S. (ed.) *Cognitive Approaches to Pedagogical Grammar*. Berlin: Mouton de Gruyter, pp. 155-196.
- Manacorda de Rosetti, M. (1961) *La gramática estructural en la escuela secundaria*. Buenos Aires: Kapelusz.
- Mańczak, W. (1980) 'Irregular sound change due to frequency in Latin', *Language Sciences*, 2, pp. 62-68.
- Manney, L. (2000) *Middle Voice in Modern Greek. Meaning and function of an inflectional category*. (Studies in Language Companion Series, Vol. 48). Amsterdam/Philadelphia: John Benjamins.
- Manning, C. D. (2003) 'Probabilistic Syntax', in Bod, R., Hay, J., and Jannedy, S. (eds.) *Probabilistic Linguistics*. Cambridge: MIT Press, pp. 289-341.
- Manzini, M. R. (1983) 'On Control and Control Theory', *Linguistic Inquiry*, 14, pp. 421-446.
- Manzini, M. R. (1986) 'On Italian si', in Borer, H. (ed.) *Syntax and Semantics 19: The syntax of pronominal clitics*. New York: Academic Press, pp. 241-262.

- Manzini, M. R., and Savoia, L. M. (2004) 'Clitics: co-occurrence and mutual exclusion patterns', in Rizzi, L. (ed.) *The Cartography of Syntactic Structures. Vol II. The Structure of CP and IP*. Oxford: Oxford University Press, pp. 211-250.
- Manzini, M. R., and Savoia, L. M. (2005) *I dialetti italiani e romanci. Morfosintassi generativa, 3 volumes*. Alessandria: Edizioni dell'Orso.
- Manzini, M. R., and Savoia, L. M. (2007) *A unification of morphology and syntax : investigations into romance and Albanian dialects*. London; New York: Routledge.
- Manzini, R., and Savoia, L. (2013) *From Latin to Romance: case loss and preservation in pronominal systems*. Available at: <http://ling.auf.net/lingbuzz/001873> (Accessed: 16th July 2016).
- Manzini, R., and Savoia, L. (2015) *Enclisis/proclisis alternations in Romance: allomorphies and (re)ordering*. Available at: <http://ling.auf.net/lingbuzz/002679> (Accessed: 16th July 2016).
- Marantz, A. (1984) *On the nature of grammatical relations*. Cambridge, MA: The MIT Press.
- Marcus, G. F. (1999) 'Language acquisition in the absence of explicit negative evidence: can simple recurrent networks obviate the need for domain-specific learning devices?', *Cognition*, 73, pp. 293-296.
- Marelj, M., and Reuland, E. (2013) 'Clitic SE in Romance and Slavonic revisited', in Kor Chahine, I. (ed.) *Current Studies in Slavic Languages*. Amsterdam: Benjamins, pp. 75-88.
- Marinucci, M. (1996) *La lingua italiana*. Milano: Mondadori.
- Martín Zorraquino, M. (1979) *Las construcciones pronominales en español : paradigma y desviaciones*. Madrid: Gredos.
- Martín Zorraquino, M. (1994) 'La estructura del predicado y los valores de se en la Gramática de la lengua española de Alarcos', *Español Actual*, 61, pp. 53-59.
- Martin, F. J. (2012) *Deconstructing Catalan Object Clitics*. Doctoral Thesis. New York University.
- Martínez, J. A. (1981) 'Sobre el se en español', *Revista de Filología de la Universidad de La Laguna*, pp. 87-92.
- Mascaró, J. (1986) *Morfologia*. Barcelona: Enciclopèdia Catalana.
- Mascaró, J., and Rigau, G. (2002) 'Introduction The Grammar of Clitics', *Catalan Journal of Linguistics*, 1, pp. 9-15.
- Masini, F. (2008) 'Costruzioni verbo-pronominali "intensive" in italiano', *XLII Congresso SLI 25-27 settembre 2008 – Scuola Normale Superiore, Pisa*. Available at: http://francescamasini.caissa.it/Presentations_files/pisa_handout_def.pdf (Accessed: 1st January 2012).
- Maslova, E. (2008) 'Reflexive Encoding of Reciprocity: Cross-Linguistic and Language Internal Variation', in Gast, V., and König, E. (eds.) *Reciprocals and Reflexives: Cross-Linguistic and Theoretical Explorations*. Berlin: Mouton de Gruyter, pp. 225-257.
- Masullo, P. (1992) *Incorporation and Case Theory in Spanish. A Crosslinguistics Perspective*. Unpublished Doctoral Thesis. University of Washington.
- Mateu, J. (2000) 'Syntactically-based lexical decomposition: The case of climb revisited', *Paper presented at the Berkeley Linguistics Society*. University of California at Berkeley, Berkeley, CA: To appear in: Proceedings of BLS 26.

- Matthews, P. H. (1979) *Generative grammar and linguistic competence*. London; Boston: G. Allen & Unwin.
- McGinnis, M. (2004) 'Lethal ambiguity', *Linguistic Inquiry*, 35, pp. 47-95.
- McIntyre, A. (2006) 'The Interpretation of German datives and English have', in Hole, D., Meinunger, A., and Abraham, W. (eds.) *Datives and Other Cases*. Amsterdam: John Benjamins, pp. 185-211.
- McKoon, G., and Macfarland, T. (2000) 'Externally and internally caused change of state verbs', *Language*, 76, pp. 833-858.
- McNally, L. (1992) *An Interpretation for the English Existential Construction*. Doctoral Thesis. University of California.
- McNally, L. (2011) 'Existential sentences', in Maienborn, C., von Stechow, K., and Portner, P. (eds.) *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter, pp. 1829-1848.
- Melander, J. (1929) 'L'origine de l'Italien me ne, me lo, te la, etc', *Studia Neophilologica*, 2, pp. 169-203.
- Melillo, M. (1981) *I pronomi dei dialetti di Puglia nelle versioni della parabola del figliol prodigo*. Università di Bari.
- Melis, L. (2004) 'Cas, fonctions et rôles – des composantes nécessaires pour décrire les constructions verbales? Une enquête préliminaire des constructions verbales en français', in Kailuweit, R., and Humme, M. (eds.) *Semantische rollen*. Tübingen: Niemeyer, pp. 168-183.
- Mendikoetxea, A. (1992) *On the nature of Agreement: The Syntax of ARB SE in Spanish*. Doctoral Thesis. University of York.
- Mendikoetxea, A. (1998) 'Aspectos semánticos y sintácticos de las oraciones medias', *Comunicación presentada en el 28º Congreso de la Sociedad Española de lingüística*.
- Mendikoetxea, A. (1999) 'Construcciones con se: medias, pasivas e impersonales', in Bosque, I., and Demonte, V. (eds.) *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, Chapter 26.
- Mendikoetxea, A. (2000) 'Relaciones de Interficie: los Verbos de Cambio de Estado', in Bravo, A., Luján, C., and Pérez, I. (eds.) *Cuadernos de Lingüística VII*. Instituto Universitario Ortega y Gasset, pp. 125-144.
- Meneghin, C. (2008) 'Rebuilding the Rhaeto-Cisalpine Written Language: Guidelines and Criteria. Part II. Morphology, I: Noun, Article and Personal Pronoun', *Ianua. Revista Philologica Romanica*, 8, pp. 113-52.
- Menéndez Pidal, R. (1904) *Manual de gramática histórica española*. 12th edition. Madrid: Espasa-Calpe.
- Mensching, G., and Remberger, E.-M. (2006) 'Studies on Agreement', in Costa, J., and Figueiredo, M. C. (eds.) *Probes: Lack of Agreement in Romance*. (= Linguistics Today, 86). Amsterdam: Benjamins, pp. 173-201.
- Mereu, L. (2011) 'Syntactic and argument structure of existentials', *Talk given at Manchester Symposium on Existentials, University of Manchester: June 2011*.
- Mertens, P. (1993) 'Accentuation, intonation et morphosyntaxe', *Travaux de Linguistique*, 26, pp. 21-69.
- Mertens, P., Goldman, J.-P., Wehril, E., and Gaudinat, A. (2001) 'La synthèse de l'intonation à partir de structures syntaxiques riches', *Traitement Automatique des Langues*, 42.1, pp. 145-192.

- Meyer-Lübke, W. (1899) *Grammatik der romanischen Sprachen* 3. Leipzig: Reisland.
- Miller, P. H., and Sag, I. A. (1997) 'French Clitic Movement without Clitics or Movement', *Natural Language & Linguistic Theory*, 15.3, pp. 573-639.
- Milsark, G. (1974) *Existential Sentences in English*. Doctoral Thesis. Massachusetts Institute of Technology.
- Mintz, T. H., Newport, E. L., and Bever, T. G. (2002) 'The distributional structure of grammatical categories in speech to young children', *Cognitive Science*, 26, pp. 393-424.
- Miró, J. (2007, *in press*) 'Clitiques, grammaires et normalisation en gascon', *Actes du VIIIe Congrès de l'AIEO (Association Internationale d'Etudes Occitanes)*. Pessac: Presses Universitaires de Bordeaux.
- Moliner, M. (1967) *Diccionario de uso del español*. Madrid: Gredos.
- Monachesi, P. (1996) 'On the representation of Italian clitics', in Kleinhenz, U. (ed.) *Interfaces in Phonology*. Berlin: Akademie Verlag, pp. 83-101.
- Monachesi, P. (2005) *The verbal complex in romance a case study in grammatical interfaces*. Oxford/New York: Oxford University Press. Available at: <http://site.ebrary.com/id/10271721> (Accessed: 16th November 2012).
- Monge, F. (1955) 'Las frases pronominales de sentido impersonal en español', in Sánchez López, C. (ed.) *Las construcciones con se*. Madrid: Visor, pp. 341-391.
- Montes Giraldo, J. J. (2003) 'El "se" del español y sus problemas', *Estudios Filológicos*, 38, pp. 121-137.
- Moreno Cabrera, J. C. (1998) 'On the relationships between grammaticalization and lexicalization', in Ramat, A. G., and Hopper, P. (eds.) *The Limits of Grammaticalization*. (Typological Studies in Language 37). Amsterdam/Philadelphia: John Benjamins, pp. 211-227.
- Morin, Y.-C., and Kaye, J. D. (1982) 'The syntactic bases for French liaison', *Journal of Linguistics*, 18, pp. 291-330.
- Morin, Y.-C. (1979) 'La morphophonologie des pronoms clitiques en français populaire', *Cahiers de linguistique*, 9, pp. 1-36.
- Moro, A. (1997) *The raising of predicates : predicative noun phrases and the theory of clause structure*. Cambridge/New York: Cambridge University Press.
- Moro, A. (1998) 'Esserci et averci: Les clitiques sujets et une analyse en «petite proposition» pour avere', in Rouveret, A. (ed.) *ÊTRE et AVOIR – syntaxe, sémantique, typologie*. Saint Denis: Presses Universitaires de Vincennes, pp. 155-170.
- Mott, B. (2007) 'Chistabino (Pyrenean Aragonese)', *Journal of the International Phonetic Association*, 37.1, pp. 103-114.
- Munaro, N. (1994) 'Alcuni casi di alternanza struttura argomentale in inglese', in Borgato, G. (ed.) *Teoria del linguaggio e analisi linguistica. XX Incontro di grammatica generativa*. Padova: Unipress, pp. 341-370.
- Murray, W. S., and Liversedge, S. P. (1994) 'Referential Context Effects on Syntactic Processing', in Clifton, C., Frazier, L., and Rayner, K. (eds.) *Perspectives on sentence processing*. Lawrence Erlbaum, pp. 359-388.
- Nagore, F. (1977) *Gramatica de la lengua aragonesa*. Zaragoza: Libreria general.

- Nagore, F. (1989) *Gramática de la Lengua Aragonesa*. Zaragoza: Mira Editores.
- Napoli, D. J. (1973[1976]) *The two si's of Italian: an analysis of reflexive, inchoative, and indefinite subject sentences in modern standard Italian*. Doctoral Thesis. Georgetown University. (Reproduced by Indiana University Linguistic Club, 1976). Available at: <http://www.swarthmore.edu/faculty-donna-jo-napoli/publications-available-download> (Accessed: 18th July 2016).
- Naro, A. (1968) *History of Portuguese Passives and Impersonals*. Unpublished Doctoral Thesis. M.I.T.
- Naro, A. (1976) 'The Genesis of the Reflexive Impersonal in Portuguese: A study in syntactic change as a surface phenomenon', *Language*, 52, pp. 779-811.
- Naughton, J. (2005) *Czech. An essential grammar*. New York: Routledge.
- Navarro, Í. (2005) *Valores de le en español mexicano. Un caso de incorporación pronominal*. Masters Thesis. Universitat Autònoma de Barcelona.
- Nespor, M., and Vogel, I. (1986) *Prosodic Phonology*. Dordrecht: Foris.
- Nevins, A. (2007) 'The representation of third person and its consequences for person-case effects', *Natural Language and Linguistic Theory*, 25, pp. 273-313.
- Nevins, A., and Săvescu, O. (2010) 'An Apparent 'Number Case Constraint' in Romanian: The Role of Syncretism', in Arregi, K., Fagyal, Z., Montrul, S. A., and Tremblay, A. (eds.) *Romance Linguistics 2008 Interactions in Romance*. Amsterdam: John Benjamins, pp. 185-200. Available at: <http://discovery.ucl.ac.uk/171425/> (Accessed: 9th July 2016).
- Newmeyer, F. J. (1998) *Language form and language function*. Cambridge, Mass.: MIT Press. Available at: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=9356> (Accessed: 29th June 2016).
- Newmeyer, F. J. (2001) 'Where is Functional Explanation?', *Chicago Linguistic Society*, 37.2, pp. 99-122.
- Newmeyer, F. J. (2003) 'Grammar is grammar and usage is usage', *Language*, 79, pp. 682-707.
- Newmeyer, F. J. (2005) *Possible and probable languages : a generative perspective on linguistic typology*. New York: Oxford University Press.
- Nicol, F. (2005) 'Romance Clitic Clusters: On diachronic changes and cross-linguistic contrasts', in Heggie, L., and Ordóñez, F. (eds.) *Clitic and Affix Combinations*. Amsterdam: John Benjamins, pp. 141-197.
- Nishida, C. (1994) 'The Spanish Reflexive Clitic se as an Aspectual Class Marker', *Linguistics*, 32, pp. 425-458.
- Nocentini, A. (2003a) 'The object clitic pronoun in Italian: A functional interpretation', in Fiorentino, G. (ed.) *Romance Objects: Transitivity in Romance Languages*. (Empirical Approaches to language Typology 27). Berlin/New York: Mouton de Gruyter, pp. 105-116.
- Nocentini, A. (2003b) 'Evoluzione e struttura dei pronomi clitici in italiano', in Maraschio, N., and Poggi Salani, T. (eds.) *Italia linguistica anno mille. Italia linguistica anno duemila. Atti del XXXIV Congresso Internazionale di Studi della Società di Linguistica Italiana*. Roma: Bulzoni, pp. 273-284.
- Noyer, R. (1997) *Features, positions and affixes in autonomous morphological structure*. New York and London: Garland Publishing.

- O'Connor, M. C. (2007) 'External possession and utterance interpretation: a crosslinguistic exploration', *Linguistics*, 45.3, pp. 577-613.
- Oca, E. (1914) 'El pronombre "se" en nominativo', *Boletín de la Real Academia Española*, pp. 573-581.
- Olszyna-Marzys, Z. (1964) *Les pronoms dans les patois du Valais central; Étude syntaxique*. Berne: Francke.
- Ordóñez, F. (2002) 'Some Clitic combination in the Syntax of Romance', *Catalan journal of linguistics*, 1, pp. 201-224.
- Ordóñez, F., and Repetti, L. (2006) 'Stressed Enclitics', in Montreuil, J.-P. (ed.) *New Analyses on Romance Linguistics: Volume II: Phonetics, Phonology and Dialectology (Selected Papers from the 35th LSRL 35)*. Amsterdam/Philadelphia: John Benjamins, pp. 167-181.
- Ormazabal, J., and Romero, J. (2001) 'A brief description of some agreement restrictions', in Fernández, B. and Albizu, P. (eds.) *Kasu eta komuntaduraren gainean. On case and agreement*. Euskal Herriko Unibertsitateko Argitalpen Zerbitzua, pp. 215-241.
- Ormazabal, J., and Romero, J. (2007) 'The Object Agreement Constraint', *Natural Language & Linguistic Theory*, 25.2, pp. 315-347.
- Oroz, R. (1966) *La lengua castellana en Chile*. Santiago: Facultad de Filosofía y Educación, Universidad de Chile.
- Otero, C. P. (1968) 'El otro se', *Actas del XI Congreso Internacional de Lingüística y Filología Románicas, Madrid, CSIC*, pp. 1841-1851.
- Otero, C. P. (1986) 'Arbitrary subjects in finite clauses', in Bordelois, I., Contreras, H., and Zagana, K. (eds.) *Generative studies in Spanish syntax*. Dordrecht: Foris, pp. 81-109.
- Otero, C. P. (1999) 'Pronombres reflexivos y recíprocos', in Bosque, I., and Demonte, V. (eds.) *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe, pp. 1427-1517.
- Otero, C. P. (2002) 'Facetas de se', in Sánchez López, C. (ed.) *Las construcciones con se*. Madrid: Visor Libros, pp. 165-206.
- Parry, M. (1998) 'The reinterpretation of the reflexive in Piedmontese: "Impersonal" se constructions', *Transactions of the Philological Society*, 96, pp. 63-116.
- Partee, B., and Borschev, V. (2002) 'Genitive of negation and scope of negation in Russian existential sentences', in Toman, J. (ed.) *Annual Workshop on Formal Approaches to Slavic Linguistics: The Second Ann Arbor Meeting 2001 (FASL 10)*. Ann Arbor: Michigan Slavic Publications, pp. 181-200.
- Partee, B., and Borschev, V. (2007) 'Existential sentences, BE, and the genitive of negation in Russian', in Comorovski, I., and von Heusinger, K. (eds.) *Existence: Semantics and Syntax*. Dordrecht: Springer, pp. 147-190.
- Pederson, J. (2005) 'The Spanish impersonal se-construction. Constructional variation and change', *Constructions 1*. Available at: www.constructions-online.de (Accessed: 1st January 2016).
- Penello, N. (2004) *Capitolo di morfologia e sintassi del dialetto di Carmignano di Brenta*. Doctoral Thesis. Università di Padova.

- Penello, N. (2006) 'I clitics locative in sardo', *Quaderni di Lavoro dell'ASIt. Progetto ASIt, Padova*, 6, pp. 1-17.
- Penny, R. (2000) *Variation and change in Spanish*. Cambridge University Press.
- Penny, R. (2002) *A History of the Spanish Language*. Cambridge University Press.
- Peperkamp, S. (1996) 'On the Prosodic Representation of Clitics', in Kleinhenz, U. (ed.) *Interfaces in Phonology*. Berlin: Akademie Verlag, pp. 102-127.
- Peperkamp, S. (1997) *Prosodic Words*. The Hague: HAG.
- Perea, M. P. (2012) 'Les Combinacions De Clítics Pronominals En Els Dialectes Catalans', *Estudis Romànics [Institut d'Estudis Catalans]*, 34, pp. 99-143.
- Perlmutter, D. M. (1971) *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.
- Perlmutter, D. M. (1983) *Studies in relational grammar*. Chicago: University of Chicago Press.
- Pescarini, D. (2007) 'Types of syncretism in the clitic systems of romance', *Anuario del Seminario de Filología Vasca Julio de Urquijo: International journal of basque linguistics and philology*, 41.2, pp. 285-300.
- Pescarini, D. (2013) 'Clitic clusters in early Italo-Romance and the syntax/phonology interface', in Meklenborg Salvesen, C., and Helland, H. P. (eds.) *Challenging Clitics*. John Benjamins Publishing Company, pp.255-281.
- Pescarini, D. (2015) 'A note on Italian datives', in Grazia Busà, M., and Gesuato, S. (eds.) *Studi in onore di Alberto Mioni*. Padova: CLEUP, pp. 491-501.
- Pescarini, D. (2016, *in press*) 'The X⁰ syntax of 'dative' clitics and the make-up of clitic combinations in Gallo-Romance', *Selected proceedings from LSRL43*. Amsterdam: Benjamins, pp. 319-337.
- Pesetsky, D. (1995) *Zero Syntax. Experiencers and cascades*. Cambridge, MA: MIT Press.
- Picallo, C., and Rigau, G. (1999) 'El posesivo y las relaciones posesivas', in Bosque, I., and Demonte, V. (eds.) *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe, pp. 973-1024.
- Pierrehumbert, J. (1999) 'Formalizing Functionalism', in Darnell, Mi., Moravcsik, E., Newmeyer, F. J., Noonan, M., and Wheatley, K. (eds.) *Functionalism and Formalism in Linguistics*. Amsterdam: Benjamins, pp. 287-304.
- Pierrehumbert, J. (2003) 'Phonetic Diversity, Statistical Learning, and Acquisition of Phonology', *Language and Speech*, 46, pp. 115-154.
- Pinker, S. (1989) *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
- Pinker, S. (1998) 'Words and rules', *Lingua*, 106, pp. 219-242.
- Piñón, C. (2001) 'A finer look at the causative-inchoative alternation', in Hastings, R., Jackson, B., and Zvolenszky, Z. (eds.) *Proceedings of Semantics and Linguistic Theory XI*. Ithaca, NY: CLC Publications, Cornell University, pp. 346-364.

- Pinto, M. (1997) *Licensing and interpretation of inverted subjects in Italian*. Doctoral Thesis. Utrechts Instituut voor Linguistiek, Universiteit Utrecht.
- Pittau, M. (1972) *Grammatica del sardo – nuorese: Il più conservativo ei parlari neolatini*. Bologna: Pàtron
- Pittau, M. (1982) *Pronunzia e scrittura del sardo-logudorese*. Sassari: Libreria Dessi. Available at: <http://trove.nla.gov.au/version/22383721> (Accessed: 12th July 2016).
- Poletto, C. (1993) 'Subject clitic-verb inversion in north eastern Italian dialects', in Belletti, A. (ed.) *Syntactic theory and the dialects of Italy*. Torino: Rosenberg & Sellier, pp. 204-51.
- Poletto, C. (1999) 'The internal structure of AgrS and subject clitics in the Northern Italian dialects', in van Riemsdijk, H. (ed.) *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, pp. 581-620.
- Poletto, C. (2000) *The higher functional field : evidence from northern Italian dialects*. New York: Oxford University Press.
- Polinsky, M. (2005) 'Applicative Structure', in Haspelmath, M., Dryer, M. S., Gil, D., and Comrie, B. (eds.) *The World Atlas of Language Structures*. Oxford University Press, pp. 442-443.
- Pollock, J.-Y. (1998) 'On the syntax of subnominal clitics: Cliticization and ellipsis', *Syntax*, 1, pp. 300-330.
- Pool, M. (1990) *Papeles temáticos y relaciones sintácticas: Un estudio de argumentos verbales en español*. Doctoral Thesis. El Colegio de México.
- Popescu, A. (2000) 'The morphophonology of the Romanian clitic sequence', *Lingua*, 110, pp. 773-799.
- Post, B. (2000) *Tonal and Phrasal Structures in French Intonation*. The Hague: Thesus.
- Postal, P. M. (1990) 'French indirect object demotion', in Postal, P. M., and Joseph, B. D. (eds.) *Studies in Relational Grammar III*. Chicago: University of Chicago Press, pp. 104-200.
- Pottier, B. (1971) *Lingüística General*. Madrid: Gredos.
- Prieto, P., D'Imperio, M., and Gili, B. F. (2005) 'Pitch accent alignment in Romance: Primary and secondary associations with metrical structure', *Language and Speech*, 48, pp. 359-396.
- Puddu, N. (2010) 'Person in Reflexive Marking from an Integrated Typological-Historical Perspective', *Presentation in the Workshop Variation and Change in Argument Realization, May 27th - 30th 2010, Napoli-Capri*.
- Pujalte, M., and Zdrojewski, P. (2013) 'Procesos de transitivización en el español del Río de la Plata', in Di Tullio, A. (ed.) *El español rioplatense: aspectos gramaticales*. Buenos Aires: Eudeba, pp. 37-58.
- Pulgram, E. (1978) 'Latin-Romance habere: double function and lexical split', *Zeitschrift für romanische Philologie*, 94.1, pp. 1-8.
- Pylkkänen, L. (2002) *Introducing arguments*. Doctoral Thesis. MIT.
- Pylkkänen, L. (2008) *Introducing Arguments*. Cambridge, MA: MIT Press.
- Quicoli, C. A. (1982) 'Some issue on the Theory of Clitics', *Linguistic Analysis*, 10.3, pp. 203-273.

- Quicoli, C. A. (1984) 'Remarks on French Clitic Systems', *Linguistic Analysis*, 14.1, pp. 55-96.
- Quirk, R., and Crystal, D. (1985) *A comprehensive grammar of the English language*. London: Longman.
- Radford, A. (1977) 'Counter-Filtering Rules', *York Papers in Linguistics*, 7, pp. 7-45.
- RAE (1973) *Esbozo de una nueva gramática de la lengua española*. Real Academia Española. Madrid: Espasa Calpe.
- Ramchand, G. (2008) *Verb Meaning and the Lexicon: A First-Phase Syntax*. Cambridge: Cambridge University Press.
- Rappaport Hovav, M. (2014) 'Lexical content and context: The causative alternation in English revisited', *Lingua*, 141, pp. 8-29.
- Recuenco, P. (1992) *Conchugación y prenombrs febles de l'aragonés*. 1st edition. (229 Berbos aragoneses conchugaus). Zaragoza: Ligallo de Fablans de l'Aragonés.
- Redington, M., Chater, N., and Finch, S. (1998) 'Distributional Information: A Powerful Cue for Acquiring Syntactic Categories', *Cognitive Science*, 22, pp. 425-469.
- Reinhart, T. (2000) 'The Theta System: syntactic realization of Verbal concepts', *OTS working papers*. Available at: http://www.let.uu.nl/~tanya.reinhart/personal/Papers/pdf/Lexic_00.pdf (Accessed: 18th July 2016).
- Reinhart, T. (2002) 'The Theta System – An Overview', *Theoretical Linguistics*, 28, pp. 229-290.
- Reinhart, T., and Reuland, E. (1993) 'Reflexivity', *Linguistic Inquiry*, 24, pp. 657-720.
- Reinhart, T., and Siloni, T. (2005) 'The Lexicon-Syntax Parameter: Reflexivization and other arity operations', *Linguistic Inquiry*, 36, pp. 389-436.
- Reinhart, T., and Tal, S. (2004) 'Against the unaccusative analysis of reflexives', in Alexiadou, A., Anagnostopoulou, E., and Everaert, M. (eds.) *The Unaccusativity Puzzle: explorations of the syntax–lexicon interface*. Oxford: Oxford University Press, pp. 159-80.
- Remberger, E.-M. (2009) 'Null subjects, expletives and locatives in Sardinian', in Georg, K., and Remberger, E.-M. (eds.) *Null subjects, expletives and locatives in Romance*. (Arbeitspapiere des Fachbereichs Sprachwissenschaft 123). Konstanz: Konstanzer, pp. 231-261.
- Renzi, L. (ed.) (1988) *Grande grammatica italiana di consultazione. La frase. I sintagmi nominale e preposizionale, vol. I*. Bologna: Mulino.
- Repetti, L., and Ordóñez, F. (2011) *Clitics of Romance Languages Database*. Available at: <http://crl.railsplayground.net> (Accessed: 1st January 2016).
- Reuland, E. (2001) 'Primitives of binding', *Linguistic Inquiry*, 32.2, pp. 439-492.
- Reuland, E. (2005) 'Binding Conditions: How are they derived?', in Müller, S. (ed.) *HPSG05: Proceedings of the 12th international conference on head-driven phrase structure grammar*. Stanford: CSLI Publications, pp. 578-593.
- Řezáč, M. (2007) 'Escaping the person case constraint: Reference-set computation in the ϕ -system', *Linguistic Variation Yearbook* 6, pp. 97-138.

- Řezáč, M. (2008) 'The syntax of eccentric agreement: The person case constraint and ansolutive displacement in Basque', *Natural Language & Linguistic Theory*, 26, pp. 61-106.
- Řezáč, M. (2010) 'Ineffability through Modularity: Gaps in the French Clitic Cluster', in Baerman, M., Corbett, G. G., and Brown, D. (eds.) *Defective Paradigms: Missing Forms and What They Tell Us*. Oxford: Oxford University Press, pp. 151-180.
- Riemsdijk, H. C. van (1999) *Clitics in the Languages of Europe*. Walter de Gruyter.
- Rigau, G. (1978) "'Hi" datiu inanimat', *Els Marges*, 12, pp. 99-102.
- Rigau, G. (1982) 'Inanimate Indirect Object in Catalan', *Linguistic Inquiry*, 13, pp. 146-150.
- Rigau, G. (1994) 'Catalan Presentational Sentences and the Properties of Agr Nodes', in Cinque, G., Koster, J., Pollock, J.-Y., Rizzi, L., and Zanuttini, R. (eds.) *Paths towards universal grammar: studies in honor of Richard S. Kayne*. Washington, D.C.: Georgetown University Press, pp. 343-359.
- Rigau, G. (1997) 'Locative Sentences and Related Constructions in Catalan: ésser/haver Alternation', in Mendikoetxea, A., and Uribe-Etxebarria, M. (eds.) *Theoretical Issues at the Morphology-Syntax Interface*. Bilbao/Donostia: UPV, pp. 395-421.
- Rigau, G. (2002) 'Els complements adjunts', in Solà, J., Lloret, M.-R., Mascaró, J., and Pérez Saldanya, M. (eds.) *Gramàtica del Català contemporani*. Barcelona: Editorial Empúries, Chapter 15, pp. 2045-2110.
- Rivera-Castillo, Y. (1997) 'Proclitic groups and word order in Caribbean Spanish', in Back, J., and Motapanyane, V. (eds.) *Clitics, Pronouns and Movement*. Amsterdam: John Benjamins, pp. 237-276.
- Rivero, M. L. (1991) 'Clitic and NP climbing in Old Spanish', in Campos, H., and Martínez-Gil, F. (eds.) *Current Studies in Spanish Linguistics*. Washington, DC: Georgetown University Press, pp. 241-282.
- Rivero, M. L. (1994) 'Clause Structure and V-Movement in the Languages of the Balkans', *Natural Language & Linguistic Theory*, 12.1, pp. 63-120.
- Rivero, M. L. (2002) 'On impersonal reflexives in Romance and Slavic and Semantic variation', in Camps, J., and Wiltshire, C. (eds.) *Romance Syntax, Semantics and L2 Acquisition. Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000*. Amsterdam/Philadelphia: John Benjamins, pp. 169-195.
- Rivero, M. L. (2003) 'Reflexive clitic constructions with datives: syntax and semantics', *Formal Approaches to Slavic Linguistics*, 11, pp. 469-494.
- Rizzi, L. (1976) 'La montée du sujet, le si impersonnel et une règle de restructuration dans la syntaxe italienne', *Recherches Linguistiques*, 4, pp. 158-185.
- Rizzi, L. (1986) 'Null objects in Italian and the theory of pro', *Linguistic Inquiry*, 17, pp. 501-557.
- Rizzi, L. (1997) 'The fine structure of the left periphery', in Haegeman, L. (ed.) *Elements of Grammar*. Dordrecht, The Netherlands: Kluwer, pp. 281-337.
- Roberge, Y., and Troberg, M. (2009) 'The high applicative syntax of the dativus commodi/incommodi in Romance', *Probus*, 21, pp. 249-289.
- Roberts, I. (1991) 'The nature of subject clitics in Franco-provençal Valdôtain', in van Riemsdijk, H., and Rizzi, L. (eds.) *Clitics and their Hosts*. Tilburg: Tilburg University, pp. 303-330.

- Roberts, M., Onnis, L., and Chater, N. (2005) 'Acquisition and evolution of quasi-regular languages: two puzzles for the price of one', in Tallerman, M. (ed.) *Language origins: perspectives on evolution*. USA: Oxford University Press. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.2514&rep=rep1&type=pdf> (Accessed: 29th June 2016).
- Roca Pons, J. (1960) *Introducción a la gramática*. Barcelona: Ed. Teide.
- Roeper, T. (1987) 'Implicit arguments and the head-complement relation', *Linguistic Inquiry*, 18, pp. 267-310.
- Rohde, D. L. T., and Plaut, D. C. (1999) 'Language acquisition in the absence of explicit negative evidence: how important is starting small', *Cognition*, 72, pp. 67-109.
- Rohlfs, G. (1949) *Grammatica storica della Lingua italiana e dei suoi dialetti. Vol. II: Morphologia*. Torino: Einaudi.
- Rohlfs, G. (1966) *Grammatica storica della lingua italiana e dei suoi dialetti. Vol. I: Fonetica*. Torino: Einaudi.
- Rohlfs, G. (1968) *Grammatica storica della lingua Italiana e dei suoi dialetti. Vol. I, II, III: fonetica, morfologia, sintassi e formazione delle parole (nel cofanetto)*. Torino: Einaudi.
- Rohlfs, G. (1969) *Grammatica storica della lingua italiana e dei suoi dialetti. Vol. II*. Torino, Einaudi.
- Rohlfs, G. (1977) *Le Gascon, Etudes de philologie pyrénéenne*. 3rd edition. (Beihefte Zur Zeitschrift für Romanische Philologie 85). Tübingen: Max Niemeyer.
- Roldán, M. (1972) 'Concerning Spanish Datives and Possessives', *Language Sciences*, 21, pp. 27-32.
- Romero, J. (1997) *Construcciones de doble objeto y gramática universal*. Doctoral Thesis. Universitat Autònoma de Madrid.
- Romero, J. (2001) 'A note on case mismatch in Spanish', *Cuadernos de Lingüística del Instituto Universitario Ortega y Gasset*, 8, pp. 97-104.
- Romero, J. (2012) 'Accusative feminine datives', in Fernández, B., and Etxepare, R. (eds.) *Variation in datives: A micro-comparative perspective*. New York: OUP, pp. 283-300.
- Romieu, M., and Bianchi, A. (2005) *Gramatica de l'occitan gascon contemporanèu*. PESSAC: Presses Universitaires de Bordeaux.
- Ronjat, J. (1913) *Essai de syntaxe des parlers. Provençaux modernes*. Elsevier.
- Ronjat, J. (1930) *Grammaire historique des parlers provençaux modernes*. Marseille, Lafitte/Geneva: Slatkine reprints [1980].
- Rooryck, J. (1992) 'Romance enclitic ordering and Universal Grammar', *The Linguistic Review*, 9.3, pp. 220-250.
- Rooryck, J., and Vanden Wyngaerd, G. (1999) 'Simplex and complex reflexives in French and Dutch', in Coene, M., De Mulder, W., Dendale, P., and D'Hulst, Y. (eds.) *Traiani augusti vestigia pressa sequamur. Studia linguistica In honorem Lilianae Tasmowski*. Padova: Unipress, pp. 617-639.
- Rooth, M. (1992) 'A theory of focus interpretation', *Natural Language Semantics*, 1, pp. 75-116.

- Roselló, J. (2002) 'El SV, I: Verb i Arguments Verbals', in Solà, J., Lloret, M., Mascaró, J., and Saldanya, M. (eds.) *Gramàtica del Català Contemporani*. Editorial Empúries: Barcelona, Chapter 13.
- Rosenbach, A. (2003) 'Aspects of iconicity and economy in the choice between the s-genitive and the of-genitive in English', in Rohdenburg, G., and Mondorf, B. (eds.) *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter, pp. 379-411.
- Rouveret, A., and Vergnaud, J.-R. (1980) 'Specifying Reference to the Subject: French Causatives and Conditions on Representations', *Linguistic Inquiry*, 11.1, pp. 97-202.
- Russi, C. (2006) 'Italian volerci: Lexical verb or functional head?', in Nishida, C. and Montreuil, J.-P. (eds.) *New Perspectives on Romance Linguistics. Vol. 1: Morphology, Syntax, Semantics, and Pragmatics*. Amsterdam/Philadelphia: John Benjamins, pp. 247-261.
- Russi, C. (2008) *Italian clitics an empirical study*. New York: Mouton de Gruyter. Available at: <http://site.ebrary.com/id/10256619> (Accessed: 27th October 2012).
- Ruwet, N. (1972a) 'Les constructions pronominales en français. Restrictions de sélection, transformations et règles de redondances', *Le Français Moderne*, 40, pp. 102-25.
- Ruwet, N. (1972b) *Théorie syntaxique et syntaxe du français*. Paris: Éditions du seuils.
- Sabatini, F. (1985) 'L"italiano dell'uso medio": una realtà tra le varietà linguistiche italiane', in Holtus, G., and Radtke, E. (eds.) *Gesprochenes Italienisch in Geschichte und Gegenwart*. Tübingen: Narr., pp. 154-184.
- Saccon, G. (1993) *Post-verbal Subjects: A Study Based on Italian and Its Dialects*. Doctoral Thesis. Harvard University.
- Sadock, J. M. (1995) 'Multi-hierarchy view of clitics', *Papers from the Parasession on Clitics*, pp. 258-279.
- Sag, I. A., and Wasow, T. (1999) *Syntactic Theory*. Stanford: CSLI.
- Sala-Gallini, M. (1996) 'Lo statuto del clítico nella dislocazione a destra: pronome vero o marca flessionale?', *Archivio Glottologico Italiano*, 81.1, pp. 76-94.
- Salvi, G. (2001) 'La nascita dei clitici romanzi', *Romanische Forschungen*, 113.3, pp. 285-319.
- Sampson, G. (2001) *Empirical Linguistics*. London: Continuum.
- Sánchez López, C. (2002) 'Las construcciones con se. Estado de la cuestión', *Las construcciones con se*. Madrid: Visor Libros, pp. 13-163.
- Sánchez López, C. (2007) 'The possessive dative and the syntax of affected arguments', *Cuadernos de Lingüística del I. U. I. Ortega y Gasset* 2007, 14, pp. 153-173.
- Sanford, A. J. (1999) 'Word meaning and discourse processing: A tutorial review', in Garrod, S., and Pickering, M. J. (eds.) *Language Processing*. East Sussex: Psychology Press, pp. 301-334.
- Sanz, M. (2000) *Events and predication a new approach to syntactic processing in English and Spanish*. Amsterdam/Philadelphia, PA: John Benjamins Pub. Co. Available at: <http://site.ebrary.com/id/10475928> (Accessed: 6th November 2014).
- Sanz, M., and Laka, I. (2002) 'Oraciones transitivas con se: el modo de acción en la sintaxis', in Sánchez López, C. (ed.) *Las construcciones con se*. Madrid: Visor Libros, pp. 311-341.

- Sardou, A.-L. (1978) *Grammaire de l'idiome niçois*. Marseille: Lafitte reprints.
- Saussure, F. (1916) *Course in General Linguistics*. Translation from the French by Wade Baskin. London: Fontana/Collins.
- Schäfer, F. (2008a) *The Syntax of (Anti-)Causatives. External arguments in change-of state contexts*. Amsterdam: John Benjamins.
- Schäfer, F. (2008b) 'Middles as Voiced Anticausatives', in Efner, E., and Walkow, M. (eds.) *Proceedings of NELS 37*. Amherst: GLSA, pp. 183-96.
- Schäfer, F. (2012a) 'The passive of reflexive verbs and its implications for theories of binding and case', *Journal of Comparative Germanic Linguistics*, 15.3, pp. 213-268.
- Schäfer, F. (2012b) 'Local Case, Cyclic Agree and the Syntax of truly Ergative Verbs', in Alexiadou, A., Kiss, T., and Müller, G. (eds.) *Local Modelling of Non-Local Dependencies in Syntax*. (Linguistische Arbeiten 547). Berlin: de Gruyter, pp. 273-304.
- Schäfer, F. (2015a) 'Medio-passives within a formal typology of Voice', *Paper presented at Roots IV, New York University*.
- Schäfer, F., and Vivanco, M. (2015) 'Anticausatives are weak scalar expressions, not reflexive expressions', *Glossa: a journal of general linguistics*, 1.18, pp. 1-36.
- Schmidely, J. (1978) 'De "ge lo" a "se lo"', *Cahiers de Linguistique Hispanique Medievale*, 4, pp. 63-70.
- Scholz, B. C., and Pullum, G. K. (2007) 'Tracking the origins of transformational generative grammar', *Journal of Linguistics*, 43, pp. 701-723.
- Schutze, C. T. (1996) *The Empirical Base of Linguistics: Grammaticality Judgements and Linguistic Methodology*. University of Chicago Press.
- Schwegler, A. (1990) *Analyticity and Syntheticity: a Diachronic Perspective with Special Reference to Romance Languages*. Berlin/New York: Mouton de Gruyter.
- Seco, M. (1972) *Gramática esencial de español*. Madrid: Aquilas.
- Seco, R. (1988) *Manual de gramática española*. 11th edition. Madrid: Aguilar.
- Segura i Llopes, C. (1998) *El parlar d'Elx a estudi: aproximació a una descripció*, volume 34 of *Temes d'Elx*. Elx: Ajuntament d'Elx.
- Sepúlveda Barrios, F. (1988) *La voz pasiva en el español del siglo XVII: contribución a su estudio*. Madrid: Gredos.
- Serianni, L. (1988) *Grammatica italiana. Italiano comune e lingua letteraria. Suoni, forme, costrutti, con la collaborazione di A. Castelvechi*. Torino: UTET.
- Seuren, P. (1976) 'Clitic pronoun clusters', *Italian linguistics*, 2, pp. 17-36.
- Seuren, P. (2009) 'The clitics mechanism in French and Italian', *Probus*, 21, pp. 83-142.
- Sheehan, M. (2006) *The EPP and Null Subjects in Romance*. Doctoral Thesis. University of Newcastle.

- Sheehan, M. (2010) ‘“Free” inversion in Romance and the Null Subject Parameter’, in Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M. (eds.) *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, pp. 231-262.
- Shibatani, M. (1994) ‘An integrational approach to possessor raising, ethical datives, and adversative passives’, *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society (=BLS 20)*, Berkeley, Berkeley Linguistics Society, pp. 461-486.
- Shlonsky, U. (1997) *Clause structure and word order in Hebrew and Arabic*. Oxford: Oxford University Press.
- Siewierska, A. (1984) *The passive: A comparative linguistic analysis*. London: Croom Helm.
- Sihler, A. L. (1995) *New comparative grammar of Greek and Latin*. New York & Oxford: Oxford University Press.
- Siloni, T. (2008) ‘The syntax of reciprocal verbs: An overview’, in König, E., and Gast, V. (eds.) *Reciprocals and reflexives: Cross-linguistic and theoretical explorations*. Berlin: Mouton de Gruyter, pp. 451-498.
- Simone, R. (1993) ‘Stabilità e instabilità nei caratteri originali dell’italiano’, in Sobrero, A. (ed.) *Introduzione all’italiano contemporaneo. Le strutture*. Roma/Bari: Laterza, pp. 41-100.
- Sistac i Vicén, R. (1993) *El Ribagorçà a l’Alta Llitera. Els parlars de la vall de la Sosa de Peralta*. Barcelona: Institut d’Estudis Catalans.
- Skårup, P. (1986) ‘L’ordre de pronoms placés dans la zone verbale en ancien occitan’, *Studia neophilologica*, 58, pp. 85-98.
- Slobin, D. I. (2005) ‘Linguistic representation of motion events: What is signifier and what is signified?’, in Maeder, C., Fischer, O., and Herlofsky, W. (eds.) *Outside-in - Inside-out. Iconicity in Language and Literature*, 4. Amsterdam: John Benjamins, pp. 307-322.
- Smith, N. B., and Bergin, T. G. (1984) *An Old Provençal Primer*. New York: Garland.
- Smyth, H. W. (1920) *Greek Grammar*. New York: American Book Company.
- Solà-Pujols, J. (1998) ‘Clitics are inflection. Evidence from Catalan’, in Fullana, O., and Roca, F. (eds.) *Studies on the Syntax of Central Romance Languages*. Girona: Publicacions de la Universitat de Girona, pp. 213-226.
- Solstad, T. (2007a) ‘Lexical Pragmatics and Unification: the Semantics of German Causal durch’, *Research on Language and Computation*, 5, pp. 481-502.
- Solstad, T. (2007b) *Mehrdeutigkeit und Kontexteinfluss: Die Spezifikation kausaler Relationen am Beispiel von durch [Ambiguity and context: the specification of causal relations]*. Doctoral Thesis. University of Oslo.
- Sorace, A. (1992) *Lexical conditions on syntactic knowledge: Auxillary selection in native and non-native grammars of Italian*. Doctoral Thesis. Edinburgh: University of Edinburgh.
- Sorrenti, E. (in prep) *Existential Sentences and Definiteness Effects in Southern Calabrese*. Doctoral Thesis. University of Hamburg.
- Spathas, G. (2010) *Focus on Anaphora: accent placement and the syntax and semantics of anaphors*. (LOT Dissertation Series). Utrecht: LOT.
- Sperber, D., and Wilson, D. (1986) *Relevance*. Oxford: Blackwell.

- Sperber, D., and Wilson, D. (1988) *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
- Spitzer, L. (1941) 'Feminización del neutro', *Revista de Filología Hispánica*, 3, pp. 339-371.
- Sportiche, D. (1995) 'Clitic constructions', in Zaring, L., and Rooryck, J. (eds.) *Phrase Structure and the Lexicon*. Dordrecht: Kluwer, pp. 213-276.
- Sportiche, D. (2014). 'Assessing unaccusativity and reflexivity, using focus alternatives to decide what gets which theta role', *Linguistic Inquiry*, 45.2, pp. 305-321.
- Stéfanini, J. (1962) *La voix pronominale en ancien et en moyen français*. Aix-en-Provence: Ophrys.
- Stefanowitsch, A. (2003) 'Constructional semantics as a limit to grammatical alternation: The two genitives of English', in Rohdenburg, G., and Mondorf, B. (eds.) *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter, pp. 413-443.
- Stefanowitsch, A. (2007) 'Linguistics beyond grammaticality', *Corpus Linguistics and Linguistic Theory*, 3.1, pp. 57-71.
- Stefanowitsch, A. (2008) 'Negative entrenchment: A usage-based approach to negative evidence', *Cognitive Linguistics*, 19.3, pp. 513-531.
- Stefanowitsch, A., and Gries, S. T. (2003) 'Collostructions: Investigating the interaction of words and constructions', *International Journal of Corpus Linguistics*, 8, pp. 209-243.
- Stepanov, A., Fanselow, G., and Vogel, R. (eds.) (2004) *Minimality effects in syntax*. (Studies in generative grammar 70). Berlin/New York: De Gruyter.
- Stroik, T. (1992) 'Middles and Movement', *Linguistic Inquiry*, 23, pp. 127-137.
- Strozer, J. R. (1976) *Clitics in Spanish*. Doctoral Thesis. UCLA.
- Stubbs, M. (1996) *Text and corpus analysis*. Oxford: Blackwell.
- Studerus, L. (1984) 'Impersonal se + lo(s), la(s): a tentative regional profile', *Southwest Journal of Linguistics*, 7, pp. 57-68.
- Suñer, M. (1982) *Syntax and Semantics of Spanish Presentational Sentence-Types*. Washington, DC: Georgetown University Press.
- Taibo, M. (n.d.) *Las Construcciones Con Se En Una Muestra Del Corpus De Referencia Del Español Actual*. Available at: http://www.contrastiva.it/baul_contrastivo/dati/sanvicente/contrastiva/Gram%C3%A1tica%20espa%C3%B1ola/Taibo,%20construcciones%20con%20se.pdf (Accessed: 1st January 2016).
- Talmy, L. (1985) 'Force Dynamics in Language and Thought', *Papers from the Parasession on Causatives and Agentivity*. Chicago: Chicago Linguistic Society, pp. 293-337.
- Talmy, L. (2000) *Toward a cognitive semantics*. Cambridge, Mass: MIT Press.
- Tanenhaus, M. K., and Trueswell, J. C. (1995) 'Sentence Comprehension', in Miller, J. L., and Eimas, P. D. (eds.) *Speech, Language, and Communication*. New York: Academic Press, pp. 217-262.
- Tekavčić, P. (1972) *Grammatica storica dell'italiano, Volume 2: Morfosintassi*. Bologna: Il Mulino.

- Tesnière, L. (1959) *Éléments de syntaxe structurale*. Paris: C. Klincksieck.
- Teulat, R. (1976) *Momento Grammatical de l'Occitan Referentiel*. Editions occitanas.
- Thornton, A. M. (1996) 'On some phenomena of prosodic morphology in Italian: accorciamenti, hypocoristics and prosodic delimitation', *Probus*, 8, pp. 81-112.
- Thornton, A. M. (2007) 'Phénomènes de réduction en italien', in Delais-Roussarie, E., and Labrune, L. (eds.) *Des sons et des sens. Données et modèles en phonologie et en morphologie*. Paris: Lavoisier, pp. 241-268.
- Tiktin, H. (1891) *Gramatica română. Etimologia și sintaxa. Revised 3rd edition*. Bucharest: Tempo.
- Tinton, M. (1982) *Grammaire limousine*. 3rd edition. Tulle: Lemouzi.
- Tissot, J. (1865[1970]) *Le patois des Fourgs, arrondissement de Pontarlier, département du Doubs*. Geneva: Slatkine Reprints.
- Todolí, J. (1992) 'Variants dels pronoms febles de 3a persona al País Valencià: regles fonosintàctiques i morfològiques subjacents', *Zeitschrift für Katalanistik/ Revista d'estudis catalans*, 5, pp. 136-160.
- Tomasello, M. (1995) 'Language is Not an Instinct', *Cognitive Development*, 10, pp. 131-156.
- Tomasello, M. (2000) 'A Usage-Based Approach to Child Language Acquisition', *Berkeley Linguistics Society*, 26, pp. 305-319.
- Tomasello, M. (2003) 'The Key is Social Cognition', in Gentner, D., and Goldin-Meadow, S. (eds.) *Language in Mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press, pp. 47-57.
- Tomasello, M. (2005) *Constructing a language*. London: Harvard University Press.
- Torres Oliva, M. (2014) *Estudio descriptivo del sistema de pronombres átonos en aragonés*. Barcelona: Universitat Pompeu Fabra.
- Torres-Tamarit, F. (2010) *Stress shift under encliticization in Formenteran Catalan: A case of an unmarked stress pattern*. Masters Thesis. Universitat Autònoma de Barcelona.
- Tortora, C. (1997) *The Syntax and Semantics of the Weak Locative*. Doctoral Thesis. University of Delaware.
- Tortora, C. (2001) 'Evidence for a null locative in Italian', in Cinque, G., and Salvi, G. (eds.) *Current Studies in Italian Syntax: Studies Offered to Lorenzo Renzi*. London: Elsevier, pp. 313-326.
- Tortora, C. (2002) 'Romance Enclisis, Prepositions, and Aspect', *Natural Language & Linguistic Theory*, 20.4, pp. 725-758.
- Traugott, E. C., and Heine, B. (eds.) (1991) *Approaches to Grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
- Traugott, E. C., and König, E. (1991) 'The semantics-pragmatics of grammaticalization revisited', in Traugott, E. C., and Heine, B. (eds.) *Approaches to Grammaticalization*. Amsterdam/Philadelphia: John Benjamins, pp. 189-218.
- Treviño, E. (1992) 'Subjects in Spanish Causative Constructions', in Hirschbühler, P., and Koerner, K. (eds.) *Romance Languages and Modern Linguistic Theory*. Amsterdam: John Benjamins, pp. 309-24.

- Tuggy, D. (1985) *Imagic Meaning and Dative Possessors in Spanish*. Available at: <http://www-01.sil.org/~tuggyd/SpanDatives/Imagic.htm> (Accessed: 1st January 2016).
- Tyler, L. K., and Marslen-Wilson, W. D. (1977) 'The On-Line Effects of Semantic Context on Syntactic Processing', *Journal of Verbal Learning and Verbal Behavior*, 16, pp. 683-692.
- Uriagereka, J. (1995) 'Aspects of the Syntax of Clitic Placement in Western Romance', *Linguistic Inquiry*, 26.1, pp. 79-123.
- Vaissière, J. (1983) 'Language-independent prosodic features', in Cutler, A., and Ladd, R. (eds.) *Prosody: Models and Measurements*. Berlin: Springer Verlag, pp. 53-66.
- Vaissière, J. (1997) 'Langues, prosodies et syntaxe', *Traitement Automatique des Langues*, 38.1, pp. 53-82.
- Van Heusinger, K. (2004) 'Focus particles, sentence meaning, and discourse structure', in Abraham, W., and Alice ter Meulen (eds.) *The composition of meaning: From lexeme to discourse*. Amsterdam: John Benjamins, pp. 167-193.
- Van Hoecke, W. (1996) 'The Latin dative', in Van Belle, W., and Van Langendoncke, W. (eds.) *The dative*. Amsterdam: John Benjamins, pp. 3-37.
- Van Valin, R. D., and La Polla, R. J. (1997) *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
- Vanelli, L. (1985) 'Typologie des pronoms sujets dans les langues romanes', in Bouvier, J.-C. (ed.) *Linguistique Descriptive: Phonétique, Morphologie et Lexique*. Aix-en-Provence: Université de Provence, pp. 161-176.
- Vázquez Rozas, V. (2006) 'Gustar-Type Verbs', in Clements, J. C., and Jiyoung, Y. (eds.) *Functional Approaches to Spanish Syntax. Lexical Semantics, Discourse and Transitivity*. Palgrave Macmillan, pp. 80-114.
- Vázquez, V. (1995) *El complemento indirecto en español*. Santiago de Compostela: Universidad de Santiago de Compostela.
- Veado, R. M. A. (1982) *Comportamento Lingüístico do Dialeto Rural-MG*. Belo Horizonte: Universidade Federal de Minas Gerais.
- Vera Luján, A. (1990) *Las construcciones pronominales pasivas e impersonales en español*. Murcia: Universidad de Murcia.
- Vergnaud, J.-R., and Zubizarreta, M.-L. (1992) 'The Definite Determiner and the Inalienable Constructions in French and English', *Linguistic Inquiry*, 23, pp. 595-652.
- Verstraete, J. (2005) 'Review of Frequency and the emergence of linguistic structure, ed. by Joan Bybee and Paul Hopper', *Language*, 81, pp. 498-501.
- Viaplana, J. (1980) 'Algunes consideracions sobre les formes pronominals clíiques del barceloní', *Filologia romànica*, 6, pp. 459-471.
- Vincent, N., and Harris, M. (1982) *Studies in the Romance Verb*. London: Croom Helm.
- Viviani, A. (2006) 'I verbi procomplementari tra grammatica e lessicografia', *Studi di grammatica italiana (a cura dell'accademia della crusca)*, 25, pp. 255-322.

- Vogel, I. (2009) 'The status of the Clitic Group', in Grijzenhout, J., and Kabak, B. (eds.) *Phonological Domains: Universals and Deviations*. Walter de Gruyter, pp. 15-46.
- Voulard, P. (1988) *Roudanen, Maritime, Gavouot, Nissart, Se Parlaves provençau*. Nice: CNDP CRDP.
- Wagner, M. L. (1941) *Fonetica storica del sardo*. Cagliari: Gianni Trois.
- Wanner, D. (1974) 'The evolution of Romance clitic order', in Campbell, R., Goldin, M., and Wang, M. (eds.) *Linguistic Studies in Romance Language*. Washington, DC: Georgetown University Press, pp. 158-177.
- Wanner, D. (1977) 'On the order of clitics in Italian', *Lingua*, 43, pp. 101-128.
- Wanner, D. (1994) 'El orden de los clíticos agrupados en castellano', *Thesaurus: Boletín del Instituto Caro y Cuervo*, 49.1, pp. 1-57.
- Wasow, T. (1997) 'Remarks on grammatical weight', *Language Variation and Change*, 9, pp. 81-105.
- Wasow, T., and Arnold, J. (2003) 'Post-verbal constituent ordering in English', in Rohdenburg, G., and Mondorf, B. (eds.) *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter, pp. 119-154.
- Wasow, T., and Arnold, J. (2005) 'Intuitions in linguistic argumentation', *Lingua*, 115.11, pp. 1481-1496.
- Welby, P. S. (2003) *The Slaying Of Lady Mondegreen, Being A Study Of French Tonal Association And Alignment And Their Role In Speech Segmentation*. Doctoral Thesis. Ohio State University.
- Whitley, M. S. (1998) 'Psych Verbs: Transitivity Adrift', *Hispanic Linguistics*, 10, pp. 115-53.
- Williams, E. (1985) 'PRO and Subject of NP', *Natural Language and Linguistic Theory*, 3, pp. 297-315.
- Williams, E. (1994) *Thematic structure in syntax*. Cambridge, MA. [etc.]: The MIT press.
- Woodcock, E. C. (1959) *A new Latin syntax*. London: Methuen.
- Wray, A. (2002) *Formulaic language and the lexicon*. Cambridge/New York: Cambridge University Press.
- Wright, E. L. (1976) 'Arbitrariness and Motivation: a New Theory', *Foundations of Language*, 14, pp. 505-523.
- Wright, S. K. (2002) 'Transitivity and change of state verbs', in Larsen, J., and Paster, M. (eds.) *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, pp. 339-350.
- Wurmbrand, S. (2014) 'Tense and aspect in English infinitives', *Linguistic Inquiry*, 45.3, pp. 403-447.
- Yamaguchi, T. (1998) *Lexical Semantic Analysis of the Causative-Inchoative Alternation in Japanese*. (=Essex Graduate Student Papers in Language and Linguistics (II), ed. by Doug Arnold.) Essex, UK: University of Essex.
- Zagona, K. T. (2002) *The syntax of Spanish*. New York: Cambridge University Press.
- Zamparelli, R. (1998) 'A theory of kinds, partitives and of/z possessives', in Alexiadou, A., and Wilder, C. (eds.) *Possessors, Predicates and Movement in the Determiner Phrase*. Amsterdam/Philadelphia: John Benjamins, pp. 261-301.
- Zamparelli, R. (2000) *Layers in the Determiner Phrase*. Doctoral Thesis. University of Rochester.

- Zanuttini, R. (1997) *Negation and Clausal Structure : A Comparative Study of Romance Languages*. New York: Oxford University Press.
- Zribi-Hertz, A. (1982) 'La construction semoyen du français et son status dans le triangle: moyen, passif, réfléchi', *Linguisticae Investigationes*, 6.2, pp. 345-401.
- Zribi-Hertz, A. (1987) 'L'Ergativité réflexive en français moderne', *Le français moderne*, 55, pp. 23-54.
- Zwicky, A. M. (1977) *On Clitics*. Indiana Bloomington: University Linguistics Club.
- Zwicky, A. M. (1994) 'What is a clitic?', in Nevis, J. A., Joseph, B. D., Wanner, D., and Zwicky, A. M. (eds.) *A Comprehensive Bibliography, 1892– 1991*. Amsterdam/Philadelphia: John Benjamins, pp. xii-xx.